A mutation-independent AAV-based gene therapy approach to treat Phenylketonuria

A thesis submitted to fulfil the requirements of the degree of Master of Philosophy

Caitlin Wanda Lucas

Children's Medical Research Institute

Faculty of Medicine and Health

The University of Sydney

2026

Declaration

This is to certify that the content of this thesis is my own work. This thesis has not been submitted for any other degree or purpose.

I certify that the intellectual content of this thesis is the product of my own work, and that all assistance received in preparing this thesis and all sources have been acknowledged.

Caitlin Wanda Lucas

`

Author Attribution Statement

This is to certify that the content of this thesis is my own work. This thesis has not been submitted for any other degree or purpose.

I certify that the intellectual content of this thesis is the product of my own work, and that all assistance received in preparing this thesis and all sources have been acknowledged. Special contributions from others are acknowledged as follows: Cindy Zhu performed flow cytometry analyses; Dr Md Musfizur Hassan and Dr Mark Graham optimised and performed LC-MS/MS analysis; Tiffany Wotton from Newborn Screening at The Children's Hospital in Westmead performed LC-MS analysis on dried blood spots.

All remaining laboratory work, data analysis and chapter drafting were carried out by the candidate. Supervisors provided guidance and feedback on experimental design and interpretation, and reviewed drafts of the thesis.

The student used ChatGPT for the purposes of text enhancement. The use of this generative AI tool includes spelling corrections, minor sentence restructuring, and clarity enhancement. The author confirms that where text was modified by generative AI, the content was reviewed for errors, inaccuracies, and bias. The author takes full responsibility for the submitted thesis, confirms the work is their own, and has used generative AI in accordance with university guidelines and policies.

Caitlin Wanda Lucas

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Samantha Ginn

Acknowledgements

I would like to express my gratitude to Associate Professor Samantha Ginn and Professor Ian Alexander for their supervision and guidance throughout this project. To Sam in particular, I am especially thankful for the opportunity to join your group and for your unwavering support over the last six years, first as an employee and later as student. This journey has not been without its challenges, both personally and academically, and I know it has not been straightforward. Your willingness to see me through difficult times gave me the chance to complete this work when I was not certain I would. I am proud of what I have achieved here, but I know it would not have been possible without your support and patience.

To Ian, thank you for your thoughtful insights, encouragement, and for always pushing me to consider the bigger picture.

To the CMRI Bioresources team, Shelley Dimech, Vanessa Scott, Bella Perakis, Mehtap Baserdem and Megan Reddel, thank you for taking care of my research animals and for being wonderfully supportive. A special shout out to Shelley for helping me import my mice and for teaching me many daunting mouse techniques.

I wish to acknowledge Dr Md Musfizur Hassan and Dr Mark Graham from Biochemical Proteomics at the Children's Medical Research Institute who helped with the development of the LC-MS/MS to look at PAH activity *in vitro*, going the extra mile to try and get the assay up and running for my submission. Thank you also to Tiffany Wotton, Dr Enzo Raineri and Natalie Brunmayer from Newborn Screening at The Children's Hospital in Westmead for performing dried blood spot analysis via LC-MS.

A huge thank you to fellow students (now or soon to be Doctors) Sharntie Christina, Lara Graves, Kate Mullany and Eva van Dijk who have been with me throughout my entire journey at CMRI. Thank you for being amazing, hardworking people and for sharing both the challenges and successes of research. It is a period of my life I will never forget.

I also thank the members past and present of the Gene Therapy Research Unit that haven't already been mentioned: Mawj Adam, Otillia Chan, Anais Amaya Colina, Sharon Cunningham, Emma Doroudian, Bradley Hall, Ann-Maree Huang, Claus Hallworth, Imogen Astruc, Neeta Khandekar, Margot Latham, Grant Logan, Jessica Mathieu, Natsuki Sasaki, Lakshmy Viswanath, Fiona Yang, and Vasilisa Zvyagina. Thank you for creating such a supportive and enjoyable atmosphere in the lab. It has been a real pleasure getting to know you all. See you on Strava.

To my family, for their unwavering support and belief in me - even if you still think my job is "giving mice cancer".

To Jacinta, who has been more like a sister than a housemate, thank you for all your unfailing support and friendship throughout this. I really could not have done this without you.

To Joe, Kerry, Lou and the rest of the team at The Noble Hops, thank you for being my second family and reminding me that I am more than my research. The support you gave me, both financial and personal, was invaluable.

Finally, I am deeply grateful for the privilege of access to higher education. This research was supported by the Australian Research Training Program (RTP) Fee Offset, the NSW Health Gene Therapy Scholarship, the Babak Shahidi Memorial Scholarship, the Esma May Scholarship, the Githa Conolly Bequest, the Children's Medical Research Institute top-up and later by direct support from the Gene Therapy Research Unit.

`

Presentations arising from this work

Poster presentation:

Lucas C, Tiffany, W, Enzo Ranieri, Ginn S, Alexander IE. rAAV-mediated CRISPR/Cas9 intronic gene editing to treat Phenylketonuria (PKU). Australasian Gene and Cell therapy Society, 13th Biennial Meeting, November 2024, Sydney.

`

Abstract

Phenylketonuria (PKU) is an autosomal recessive disorder of phenylalanine metabolism, primarily caused by mutations in the phenylalanine hydroxylase (PAH) gene, resulting in the toxic accumulation of phenylalanine. Affected infants are born healthy, but without intervention, can rapidly experience cognitive and developmental delays. Although newborn screening and early dietary intervention prevent neurological damage, long-term dietary adherence is burdensome, and existing pharmacological interventions benefit only a subset of patients. Consequently, there is a real need for broadly applicable, long-term treatment options. Gene therapy holds promise for addressing this unmet need. This thesis aimed to introduce a functional copy of human PAH cDNA into the first intron of the endogenous Pah locus in a murine PKU model using a homology-independent targeted integration (HITI) approach. A dual AAV CRISPR/SaCas9 and HITI donor strategy incorporating a 2A element was developed to enable bicistronic expression under endogenous promoter control, providing a mutation-agnostic approach capable of treating all PKU variants. Proof-of-concept studies with GFP donor demonstrated efficient genomic targeting, robust expression in up to 70% of hepatocytes, and production of hybrid transcripts, confirming on-target integration. Therapeutic PAH donors exhibited similar integration and expression, although significant reductions in blood phenylalanine levels were not observed. These findings highlight both the promise and the challenges of gene editing in treating PKU. While targeted integration at the Pah locus is feasible, translation into meaningful metabolic correction will require improvements in nuclease and donor delivery. Emerging dual-modality approaches, such as using lipid nanoparticle-mediated Cas9 delivery with AAV-donor, may enhance safety and efficiency. Overall, this work supports the potential of mutation-agnostic genome editing for treating PKU.

`

Table of Abbreviations

AAV	Adeno-associated virus
AdV	Adenovirus
Alb	Albumin
APAP	Acetaminophen
ATTR	Transthyretin amyloidosis
BGHpA	Bovine growth hormone polyadenylation signal
BH4	Tetrahydrobiopterin
bp	Base pairs
CAR-T	Chimeric Antigen Receptor-T
cDNA	Complementary DNA
CMRI	Children’s Medical Research Institute
CMV	Human cytomegalovirus
cohPAH	Codon-optimised human PAH
CPS1	Carbamoyl phosphate synthetase 1
CRISPR	Clustered regularly interspaced short palindromic repeats
DAPI	4’, 6-Diamidine-2'-phenylindole dihydrochloride
dCas9	Dead Cas9
DMD	Duchenne Muscular Dystrophy
dPCR	Digital polymerase chain reaction
ddPCR	Digital droplet polymerase chain reaction
DMEM	Dulbecco’s Modified Eagles Medium
DMSO	Dimethyl sulfoxide
dNTP	Deoxyribonucleic acid
DSB	Double stranded breaks
EDTA	Ethylenediaminetetraacetic acid
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FRG	Fah$^{-/-}$Rag2$^{-/-}$Il2rg$^{-/-}$
EGFP	Enhanced green fluorescent protein
GLUL	Glutamine synthetase

`

HCC	Hepatocellular Carcinoma
hcoPAH	Human codon-optimised PAH cDNA
HBSS	Hank’s Balanced Salt Solution
HDR	Homology-independent target integration
HITI	Homology-independent targeted integration
HR	Homologous recombination
IEM	Inborn errors of metabolism
IMDM	Iscove’s Modified Dulbecco’s Medium
iPSCs	Induced pluripotent stem cells
Indels	Insertions/Deletions
IRES	Internal ribosome entry site
ITRs	Inverted terminal repeats
I.V	Intravenous
JAX	The Jackson Laboratory
LC-MS/MS	Liquid chromatogram tandem mass spectrometry
LNP	Lipid nanoparticle
MFI	Mean fluorescence intensity
mRNA	Messenger RNA
MSDS	Material safety data sheet
NBS	Newborn screening
NHEJ	Non-homologous end joining
NHP	Non-human primates
NTBC	2-(2-nitro-4-trifluoromethylbenzoyl)-1,3- cyclohexanedione
OTC	Ornithine transcarbamylase
PAH	Phenylalanine hydroxylase
Pah^{enu2}	PKU mouse model
PAL	Phenylalanine ammonia lyase
PAM	Protospacer adjacent motif
PBS	Phosphate buffered saline
PCR	Polymerase chain reactions
PEG	Polyethylene glycol
PFA	Paraformaldehyde

Phe	Phenylalanine
PKU	Phenylketonuria
qPCR	Quantitative polymerase chain reaction
rAAV	Recombinant AAV
RNA	Ribonucleic acid
RT	Room temperature
SaCas9	*Staphylococcus aureus* Cas9
SCID-X1	X-linked severe combined immunodeficiency
SD	Standard deviation
sgRNA	Single guide RNA
SMA	Spinal muscular atrophy
SNPs	Single nucleotide polymorphisms
SpCas9	*Streptococcus pyogenes* Cas9
TALENs	Transcription activator-like effector nucleases
TBG	Human thyroxine binding globulin promoter
U6	RNA polymerase III promoter
VCN	Vector copy number
vg	Vector genomes
WT	Wildtype
XLMTM	X-linked myotubular myopathy
ZFNs	Zinc-finger nucleases

`

Table of Contents

Declaration 2

Author Attribution Statement 3

Acknowledgements 4

Presentations arising from this work 5

Abstract 6

Table of Abbreviations 7

List of Figures 14

List of Tables 15

1. Introduction 16

  1.1 Gene Therapy 17

  1.2 Gene Therapy Vectors 18

    1.2.1 Adenoviral Vectors 18

    1.2.2 Lentiviral vectors 19

    1.2.3 Recombinant AAV viral vectors 19

    1.2.4 21

    1.2.5 Lipid nanoparticles as non-viral vectors 24

  1.3 Liver as a therapeutic target 25

  1.5 Inborn errors of metabolism 28

  1.6 Phenylketonuria 28

  1.6.1 Epidemiology 30

  1.6.2 Phenylalanine metabolism 31

    1.6.3 Neuropathology 34

    1.6.4 Genetic basis of PKU 35

  1.7 Existing therapies for PKU 36

  1.7 Models of PKU 38

    1.7.1 *In vitro* models of PKU 39

    1.7.2 Animal models of PKU 39

    1.7.3 41

  1.8 Novel Therapeutics for PKU 45

    1.8.1 Novel drug approaches 45

    1.8.2 Cell therapy to treat PKU. 46

    1.8.3 Gene therapy to treat PKU 47

  1.9 Limitations and future directions 62

  1.10 Requirements for success in the case of PKU 62

  1.11 Hypotheses and Aims 64

2. Materials and Methods 66

2.1 Materials 66

2.1.1 Kits and reagents 66

2.1.2 Buffers and Solutions 68

2.1.3 Chemicals 70

2.1.4 Immortalised cell lines 72

2.1.5 Competent bacterial cells for cloning 72

2.1.6 Antibodies 72

2.1.7 Plasmids 72

2.1.8 PCR primers 74

2.1.9 PCR probes 75

2.2 Tissue Culture 75

2.2.1 Routine culturing and cell counting 75

2.2.2 Cell recovery from cryostorage 76

2.2.3 Cryopreservation 76

2.3 Molecular biology 77

2.3.1 DNA extraction from cells 77

2.3.2 Restriction enzyme digest 77

2.3.3 Agarose gel electrophoresis 78

2.3.4 Purification of DNA from agarose gel slices 78

2.3.5 A-tailing of PCR DNA fragments 79

2.3.6 DNA ligation 79

2.3.7 Transformation of competent cells 79

2.3.8 Sequencing of cloned products 80

2.3.9 Propagation and isolation of plasmid DNA 80

2.3.10 Nucleic acid quantification 81

2.4 Amplification of DNA fragments using polymerase chain reaction (PCR) 82

2.4.1 General PCR 82

2.4.2 Quantitative PCR (qPCR) 83

2.4.3 Digital Droplet PCR (ddPCR) 83

2.4.4 Digital PCR (dPCR) 84

2.5 rAAV vector production 84

2.5.1 Production of rAAV by calcium phosphate triple transfections 84

2.5.2 rAAV purification through CsCl gradients 85

2.5.3 Quantification of vector particles in viral fractions by qPCR 87

2.5.4 Quantification of purified vector stocks by ddPCR 88

2.6 Constructs of rAAV plasmids and *in vitro* testing 88

2.6.1 Construction of ssAAV-CRISPR/Cas9 vectors with sgRNAs 88

2.6.2 *In vitro* guide screening 89
2.6.3 Construction of scAAV HITI donor vectors 90
2.7 Murine model of PKU 91
2.7.1 Genotyping 92
2.7.2 Blood sampling 93
2.7.3 Blood phenylalanine and tyrosine quantification 94
2.7.4 Delivery of rAAV vectors to mice 94
2.7.5 High protein diet challenge protocol 95
2.7.6 Standard harvest protocol 95
2.7.7 Isolating of murine hepatocytes by perfusion 95
2.8 Liver tissue processing 98
2.8.1 Phenol-Chloroform DNA extraction from liver tissue 98
2.8.2 RNA extraction and cDNA synthesis 99
2.8.3 Liver protein extraction and quantification 100
2.8.4 Sectioning of cryopreserved tissue 100
2.9 Molecular analyses of murine liver samples 101
2.9.1 Quantification of liver tissue vector copy number using dPCR and ddPCR 101
2.9.2 Genomic DNA junction analysis using dPCR 101
2.9.3 Transcript analysis by dPCR 102
2.9.4 PAH *in vitro* activity assay 102
2.9.5 Flow cytometric analysis of GFP expression in hepatocytes 104
2.9.6 Immunofluorescent co-staining of GFP and GLUL in frozen liver sections 105
2.9.7 Microscopy and image analysis 106
2.10 Statistical analysis and figure generation 106
3. Genome editing murine *Pah* locus using CRISPR/Cas9 system 107
3.1 Introduction 107
3.2 Results 109
3.2.1 PKU mouse model colony establishment and phenotype characterisation 109
3.2.2 Mutation-independent genome editing strategies to treat PKU 111
3.2.4 Evaluation of sgRNAs for SaCas9 in murine cell lines *in vitro* 115
3.2.5 Evaluation of sgRNAs for SaCas9 in murine model of PKU *in vivo* 116
3.2.6 Design of a dual rAAV vector system to enable targeted insertion of donor-DNA vector when paired with CRISPR-Cas9 vector 118
3.2.8 Quantification of perfect HITI genomic junctions showed modest levels of targeted integration into *Pah* intron 1 126
3.2.9 Analysis of transcript in dual treated mice confirms successful on-target editing 130
3.2.10 Indel and vector copy number analyses show no difference between GFP and PAH donor vector. 134

3.2.11 Blood phenylalanine levels were not significantly reduced despite reversion of light-fur phenotype in dual vector treated *Pah*$^{enu2/enu2}$ mice. 137

3.2.12 Development of an *in vitro* PAH enzymatic assay by LC-MS/MS 140

3.3 Discussion 143

3.3.13 Conclusion 152

4. Discussion, conclusions and future perspectives 153

4.1 General discussion 153

4.2 Limitations and Future directions 154

4.3 Conclusions 158

5. References 160

6. Appendices 177

6.1 Methods and Analysis tools 177

6.1.1 pAAV2-LPhOTC BB2 backbone plasmid map 177

6.1.2 Synthesised donor DNA sequences 178

6.1.3 QuPath image analysis software GFP-positive cell counting macro 179

6.1.4 Flow cytometry gating strategy 181

6.2 Supporting Data 182

6.2.1 Phenylalanine supplementary drinking water 182

6.2.2 Additional photos of coat colour over time 183

6.2.3 All raw matched data 184

`

List of Figures

Figure 1.1: Structure of the hepatic lobule and metabolic zonation within the acinus. 27
Figure 1.2: Metabolism of phenylalanine 33
Figure 1.3: Genetic basis of PKU 37
Figure 1.4: Overview of CRISPR-Cas9 genome editing outcomes following double strand break formation (DSB) 56
Figure 3.1: Establishment of Pah^{enu2} mouse model for PKU..110
Figure 3.2: Diagram of gene-editing approaches at intron 1 of the murine Pah locus........................113
Figure 3.3: Location of selected sgRNA targets across murine Pah intron 1....................................114
Figure 3.4: Evaluation of candidate SaCas9 sgRNAs *in vivo*..117
Figure 3.5: Design and testing of mutation-independent genome editing approach using dual AAV CRISPR/SaCas9 and donor repair templates *in vivo*. 120
Figure 3.6: Analysis of GFP expression in dual and donor-only treated mice reveals GFP expression in a high proportion of hepatocytes *in vivo*. 125
Figure 3.7: Detection of HITI genomic junctions in dual-treated mice by dPCR. 129
Figure 3.8: Quantification of hybrid Pah transcripts in mice following AAV delivery.......................133
Figure 3.9: Indel and vector copy number analysis do not indicate difference in approach performance in both hybrid and universal approaches..136
Figure 3.10: Blood phenylalanine levels were not significantly reduced despite reversion of light-fur phenotype..139
Figure 3.11:Development of an *in vitro* PAH enzymatic activity assay using LC-MS/MS...............142

Appendices Figure 6.1: Flow cytometry gating strategy for GFP+ hepatocyte detection. 181
Appendices Figure 6.2: Blood amino acid levels in heterozygous and homozygous Pahenu2 mice when water supplemented with phenylalanine. 182
Appendices Figure 6.3: Representative additional images of fur phenotype reversion following dual CRISPR/Cas9 and PAH donor vector treatment. 183

`

List of Tables

Table 1.1: Animal models of PKU help elucidate pathogenesis and allow therapeutic testing 44
Table 1.2: Recent and ongoing clinical trials for PKU. 58
Table 1.3: Preclinical gene editing studies to treat PKU. 61
Table 2.1: Kits and reagents. 66
Table 2.2: Buffers and Solutions. 68
Table 2.3:Chemicals. 70
Table 2.4: Immortalised mammalian cell lines. 72
Table 2.5: Competent bacterial Escherichia coli strains used for cloning. 72
Table 2.6: Antibodies. 72
Table 2.7: Plasmids. 72
Table 2.8: PCR primers. 74
Table 2.9: PCR probes. 75
Table 2.10: Top and bottom oligonucleotides coding each gRNA sequence. 89
Table 3 1: Candidate SaCas9 sgRNAs and their respective PAM sequences†. 115
Appendices Table 6 1: Synthesised donor DNA sequences. 178
Appendices Table 6 2: Raw experimental data for Guide screening experiment. 184
Appendices Table 6 3: Raw experimental data for Cas9 + donor vector-treated cohort. 185

`

1. Introduction

Phenylketonuria (PKU) is one of the most common inherited metabolic disorders, with a worldwide incidence of 1 in 24,000 live births and an estimated 450,000 people currently affected (1). PKU is caused most frequently by mutations in the gene encoding phenylalanine hydroxylase (*PAH*), a hepatic enzyme responsible for the conversion of phenylalanine to tyrosine. I In the absence of function PAH, phenylalanine accumulates in the blood and brain, leading to severe neurotoxicity if left untreated. Newborn screening programmes and early dietary interventions have dramatically improved clinical outcomes, however the burden of strict lifelong dietary management remains substantial, and alternative pharmacological therapies are only appliable in a subset of patients. These limitations underscore the need for definitive, long-term therapies.

Gene therapy offers the potential to correct the underlying metabolic defect by restoring functional PAH activity in hepatocytes. Over the past three decades, a wide range of strategies have been explored, ranging from early viral vector–based gene addition to more recent genome editing, base editing, and prime editing technologies. Among available delivery platforms, adeno-associated virus (AAV) has emerged as the most widely used platform, with numerous preclinical studies and clinical trials highlighting both its therapeutic promise and associated challenges.

This chapter reviews the current landscape of gene therapy, outlines major delivery platforms, introduces the liver as a therapeutic target, and summarises preclinical and clinic efforts to treat PKU. It then examines the barriers that have limited clinical translation and discusses the emergence of genome editing strategies as a potential means to achieve durable therapeutic benefit.

`

1.1 Gene Therapy

Gene therapy refers to the use of genetic material to correct, replace or modulate genes for therapeutic purposes. It can be applied to the treatment of both inherited and acquired conditions, including cancer, cardiovascular diseases, and infectious disorder (2). The successful of any gene therapy depends not only on efficient delivery of genetic material to target cells, but also on achieving adequate gene expression and proper regulation of transgene activity while minimising off-target effects and immunogenicity Genetic material, most commonly DNA or mRNA, can be delivered in several approaches, including encapsulation within viral particles or nanoparticles, collectively referred to as vectors. The resulting therapeutic effect many be transient or long-lasting depending on the delivery methods and genetic cargo delivered.

Gene therapy can be broadly categorised into gene addition or gene editing. Gene addition involves supplying a function copy of a gene, which is typically maintained outside of host genome as an episome. While this can provide therapeutic benefit, the introduced genetic material is gradually lost as the cells divide, alongside any therapeutic benefit. In contrast, gene editing directly modifies host genome to elicit a durable therapeutic effect, such as correcting a pathogenic mutation or disrupting a toxic gene product. As of 2025, the U.S Food and Drug Administration (FDA) has approved 36 gene therapies, with more than 3,900 clinical trials completed or on-going worldwide (2). These therapies are delivered *in vivo*, by direct administration of the vector to the patient or *ex vivo*, where cells are genetically modified outside the body before re-infusion (3). *In vivo* gene therapies can be administered either systemically, most commonly by intravenous (IV) infusion, or locally, for example by intramuscular or intracerebral injection.

`

Several landmark approvals illustrate the transformative potential of this field. Casgevy, an ex *vivo* therapy for sickle cell disease that prevents painful vaso-occlusive crises; Luxturna, an *in vivo* treatment for RPE65-mediated inherited retinal disease that restores functional vision; Zolgensma, an *in vivo* therapy to treat spinal muscular atrophy (SMA) that enables children survive beyond 18 months, and in many cases, achieve and maintain developmental motor milestones. Collectively, these successes underscore both the life-changing impact of gene therapies.

1.2 Gene Therapy Vectors

Modern viral vector platforms include adenoviruses, lentiviruses and adeno-associated viruses, each with distinct advantages and limitations (4).

1.2.1 Adenoviral Vectors

Adenovirus (AdVs) are non-enveloped, double-stranded DNA viruses with an icosahedral capsid (~100 nm) capable of accommodating large transgene cassettes up to 36 kb. They can efficiently transduce both dividing and non-dividing cells, and the wide range of identified AdV serotypes provides broad tissue tropism (5). AdVs persist as episomes and do not typically integrate into the host genome. However, as common human pathogens responsible for respiratory, ocular, and gastrointestinal infections, most individuals possess pre-existing neutralising antibodies that can significantly limit AdV vector efficacy (6,7).

As of March 2023, AdV vectors account for approximately 15% of gene therapy clinical trials worldwide (2). However, their utility is constrained by the induction of strong innate and adaptive immune responses (8). A tragic example was the death of Jesse Gelsinger in 1999 during an Ad5-based gene therapy trial for ornithine transcarbamylase (OTC) deficiency, in

`

which overwhelming immune activation led to multiorgan failure (9). Despite these setbacks, AdVs remain clinically relevant in cases where immunological activation can contribute to therapeutic efficacy. For example, Gendicine, an AdV carrying a p53 transgene under the control of a Rous sarcoma virus promoter, was approved in China as the world's first commercialized gene therapy drug for cancer (10).

1.2.2 Lentiviral vectors

Lentiviruses are spherical, single-stranded RNA retroviruses approximately 100 nm in diameter which can transduce non-dividing cells. Lentiviral vectors can package inserts up to 9 kb (11) and integrate stability into the host genome, with a preference for transcriptionally active sites (12). This integration provides durable gene expression and has been critical for the success of *ex vivo* cell therapies, including the CAR-T cell therapy Kymriah (13) and the hematopoietic stem-cell therapies Casgevy for sickle cell disease (14) and Skysona for cerebral adrenoleukodystrophy (15). However, integration carries risks: approximately 4% of participants in early trials for Skysona developed myelodysplastic syndrome—and with extended follow-up, the overall incidence rose to about 15%—a malignancy clearly attributed to vector integration into proto-oncogenes (16). Current lentiviral designs have addressed many of these risks through improved self-inactivating (SIN) vectors and refined safety features, but insertional oncogenesis remains a theoretical concern (17). As of March 2023, lentiviral vectors account for ~10% of gene therapy clinical trials worldwide, typically used in *ex vivo* gene therapy (2).

1.2.3 Recombinant AAV viral vectors

Adeno-associated viruses (AAVs) are small, non-enveloped parvoviruses, approximately 26 nm in diameter, with an icosahedral protein capsid enclosing an approximately 4.7 kb single-

`

stranded DNA genome (18). Wildtype AAVs require co-infection with a helper virus, such as AdV, for replication. Although AAVs have been identified in multiple vertebrate species, including humans and non-human primates (NHPs), but are generally regarded as non-pathogenic (19). Natural AAV infection commonly occurs in childhood, although often asymptomatic, and leads to long lasting seroconversion, with approximately 80% of adults carrying neutralising antibodies (32). Wildtype AAV can integrate into the host genome but does so in a site-specific integration, preferentially inserting into the intronic region on chromosome 19 known as AAVS1 (20). However, recent studies have demonstrated wildtype AAV integration can also occur at alternative genomic sites, including known cancer driver genes, with AAV insertions detected in 2-5% of hepatocellular carcinoma (HCC) samples, suggesting a possible role in tumourigenesis and a more complex biology than previously assumed (21,22).

The AAV capsid is comprised of 60 viral protein subunits in a ratio of 1:1:10 (VP1:VP2:VP3) (23). Capsid structure determines tissue tropism through interactions with cellular receptors, most notably the AAV receptor (AAVR) (24). Variations in capsid sequence form the basis for AAV serotype classification, with more than a dozen of naturally occurring serotypes identified. These exhibit distinct tissue transduction profiles; for example, AAV1 preferentially transduces muscle, whereas AAV5 shows enhanced transduction of the lung following systemic injection (25). In addition to these natural serotypes, hundreds of engineered variants have been developed to alter tropism, improve transduction efficiency, and overcome pre-existing immunity (26).

The AAV genome contains two open reading frames (ORFs): *rep*, encoding non-structural proteins involved in replication and integration, and *cap*, encoding the structural capsid proteins. The genome is flanked by inverted terminal repeat sequences (ITRs), palindromic

`

elements essential for viral replication and packaging (27). In recombinant AAV (rAAV) vectors, the *rep* and *cap* genes are replaced with an expression cassette encoding a transgene of interest. During vector production, r*ep*, c*ap* and essential helper functions are supplied in trans, preventing autonomous replication and site-specific integration. As a result, rAAV genomes persist primarily as episomal concatemers within the nucleus of transduced cell (28). Although initially thought to be non-integrating, rAAV can integrate infrequently, particularly at sites of DNA damage or transcriptional activity, a process facilitated by the ITRs (29).

Most rAAV vectors are produced as single-stranded AAV (ssAAV), which requires host-cell second-strand synthesis, a recognised rate-limited step for transgene expression (30). To overcome this limitation, self-complementary AAV (scAAV) vectors were developed, in which a mutated ITR enables the genome to fold into double-stranded DNA upon entry. This enables faster and more robust transgene expression, albeit at the cost of halving the packaging capacity to approximately 2.3 kb (31,32).

Despite widespread seropositivity (33), rAAV remains an attractive vector due to its relatively low immunogenicity, ability to transduce dividing and non-dividing cells and low levels of viral integration into the genome. Nonetheless, humoral and cellular immune responses can limit efficacy and preclude both initial dosing and the ability to redosing (34). As of March 2023, AAV vectors account of ~9% of all gene therapy clinical trials (2), a proportion that continues to rise considering their favourable safety profile and therapeutic successes.

1.2.4 Challenges facing use of AAV vectors for gene therapy

While AAVs have proven to be highly effective vectors for delivering therapeutic DNA, they

`

present significant challenges. Immune responses to AAV capsids and transgenes, the high doses often required to achieve efficacy, and the potential for vector integration into the host genome constitute major safety considerations for AAV-based gene therapies. Clinical trials for both approved AAV gene therapies for haemophilia A and B, SMA, and investigation therapies such as Astellas' AT132 for X-linked myotubular myopathy (XLMTM) have reported hepatotoxicity following systemic administration, with some cases resulting in patient deaths. These events have provided a sobering reminder of the risks associated with AAV gene therapies and prompted the FDA to release a safety communication in 2021 reviewing known toxicities (35). Investigation into the systemic and cellular response to AAV gene therapies is imperative for patient safety, especially in conditions that are not fatal and thus the cost: benefit ratio is skewed. A detail overview of the immune cascade and cellular responses to rAAV vectors is beyond the scope of this chapter but is comprehensively described elsewhere (36).

1.2.4.1 Systemic response

Systemic immune response to both AAV capsid and transgene have been observed, driven by innate immune response and adaptive CD8+ cell response (37). Pre-existing neutralising antibodies, arising from natural exposure to wildtype AAV, are common (frequency in population 40-80%) (33) and often serve as exclusion criteria in clinical trials. Moreover, the formation of *de novo* neutralising antibodies post-treatment precludes vector re-administration, underpinning the current 'one-and-done' paradigm in AAV gene therapy (38). Elevation in liver transaminases, indicative of liver inflammation, have been widely reported 4-8 weeks post-vector infusion, representing delayed hepatotoxicity. These events can usually be managed with corticosteroids (37). However, severe outcomes have also been observed. Approximately one-third of patients receiving a dose of 1.1×10^{14} vg/kg of Zolgensma, an

`

AAV9 gene addition therapy for SMA, experienced hepatotoxic adverse events, and two children subsequently died of acute liver failure within 6 weeks of treatment (39,40). In the AT132 trial for XLMTM, 4 children died >1 year after infusion from progressive liver failure at high doses of vector (41).

Other toxicities observed includes HCC and thrombotic microangiopathy have also been observed in human and animal studies post gene therapy (42,43). Importantly, in non-hepatic indication, much of the infused vector is sequestered by the liver, exacerbating off-target toxicity while reducing the therapeutic delivery to the target tissue (44).

With current capsid technology and therapeutic approaches, doses of greater than 1.1×10^{14} vg/kg are often required to achieve efficacy in neuromuscular or neurological disorders such as SMA, Duchenne muscular dystrophy (DMD) and XLMTM. At these levels – representing over one quadrillion viral particles in a small child, far exceeding the number of cells in the human body—life-threatening complications have occurred, including patient deaths (38). As these disorders often already involve hepatic manifestations, AAV administration may exacerbate underlying pathology. The immunogenicity of AAV vectors, the dose required, and integration frequency are tightly linked; therefore, strategies to reduce the effective therapeutic dose are likely to improve safety.

1.2.4.2 AAV integration

While rAAV were initially believed to be non-integrating, it is now well established that rAAV genomes can integrate into the host genome a low frequence, particularly at high vector doses. Reported integration of AAV vectors post AAV gene therapy vary widely in the literature, with some studies reporting 1-3% of vector genomes randomly integration into the genome (45,46) while others suggest true random integration frequency is much lower at 0.001-0.7% (47). Recombinant AAV vector integration events are random but biased towards accessible genomic regions, occurring preferentially at regions of DNA damage, transcriptionally active loci, and open chromatin.

Although rare, rAAV integration raises important safety concerns because insertion within or near active genes can disrupt normal gene function or alter gene regulation. Integration proximal to proto-oncogenes or tumour suppressor genes may promote insertional mutagenesis and contribute to malignant transformation. A subset (2-5%) of hepatocellular carcinoma cases were found to contain AAV integration events following natural wildtype AAV infection, with sites of integration found in known cancer driver genes, suggesting a possible role in tumourigenesis (21,22). A small number of participants in AAV liver gene therapy trials, such as one cases in the recent HOPE trials for haemophilia B, developed HCC post treatment. However, this patient was already at an elevated risk for HCC, and it was determined to be unlikely related to AAV integration events found in the tumour sample (rate of 0.025% of cells in sample)(48). As a result, AAV integration and its implications on safe use as a vector remains an active area of research, with many now looking towards non-viral or genome editing based approaches to mitigate risk.

1.2.5 Lipid nanoparticles as non-viral vectors

Non-viral gene therapies are increasingly being explored as alternatives to viral vectors for gene therapy in order to mitigate limitations such as immunogenicity and packaging constraints. Among these, lipid nanoparticles (LNPs), have emerged as the most widely used platform, gaining notoriety through their successful application in COVID-19 vaccines. LNPs

`

are lipid vesicles typically composed of four key components: an ionisable cationic lipid, a helper lipid, a polyethylene glycol (PEG) lipid, and a cholesterol (49).

LNPs can efficiently transport nucleic acids and protein to the cytoplasm, however delivery to the nuclear remains limited and represents a major barrier to broader therapeutic applications. Following intravenous injection, LNPs display a natural tropism for the liver, although surface and formulation modifications can be employed to redirect delivery towards alternative target tissues (50). Compared to viral vectors, LNPs exhibit relatively low immunogenicity, enabling repeat dosing, a major advantage over viral systems.

Clinically, LNP mediated gene editing is being adopted in the development on therapies for several conditions, including haemophilia A and B and transthyretin amyloidosis (ATTR), hereditary angioedema (51–54). However, challenges remain: LNP-mediated delivery is inherently transient due to rapid degradation and clearance *in vivo*, which may limit therapeutic persistence but is advantageous for applications requiring short-lived expression. Despite these limitations, the pace of LNP development is accelerating rapidly, with a growing number of clinical trials reflecting their expanding therapeutic potential (2).

1.3 Liver as a therapeutic target

The liver is the central metabolic organ of the body, with an estimated 500 separate biochemical processes occurring in a single liver cell (55). It performs essential functions that include (i) the metabolism of amino acids, proteins, carbohydrates, hormones, lipids, and lipoproteins (ii) storage of glycogen, vitamins and minerals, (iii) detoxification and purification of blood (iv) synthesis of plasma proteins (v) production and secretion of bile, and (vi) excretion of bilirubin, cholesterol, hormones and drugs. The liver is composed primarily of hepatocytes - specialised epithelial cells that make up roughly 80% of its mass and are responsible for most of its functions.

To achieve such diverse function, the liver is organised into structural units known as lobules (56) (Figure 1.1). Lobules are typically hexagonal in shape and are centred around a central vein with portal triads (consisting of a branch of the portal vein, hepatic artery, and bile duct) positioned around its corners (Figure 1.1 A). Blood from the portal vein (nutrient-rich) and the hepatic artery (oxygen-rich) flow inwards through specialised capillaries called sinusoids, which are lined be fenestrated endothelial cells. These fenestrations permit the exchange of metabolites between the blood and the surrounding hepatocytes (57,58) (Figure 1.1 B). For gene therapy applications, the sinusoid fenestrations (100-200 nm in diameter) are particularly relevant, as they permit the diffusion of AAV viral vectors (~25 nm). Hepatocytes are arranged in plates radiating towards the central vein, where blood is collected while bile flows in the opposite direction outwards towards the bile ducts at the lobule periphery (Figure 1.1 C).

A defining feature of the lobule is metabolic zonation, whereby hepatocytes perform distinct functions depending on their position along the periportal-to-pericentral axis. Periportal hepatocytes (near the portal triads) operate in an oxygen- and nutrient-rich environment and specialise in oxidative metabolism, gluconeogenesis, and the urea cycle. In contrast, pericentral hepatocytes (near the central vein) are exposed to low oxygen and nutrient levels, and instead specialise in glycolysis, drug metabolism and ammonia detoxification via glutamine synthetase. This spatial division of labour ensures efficient handling of nutrients and toxins. Disruption of these processes leads to metabolic imbalance and contributes to the pathogenesis of many diseases (55).

`

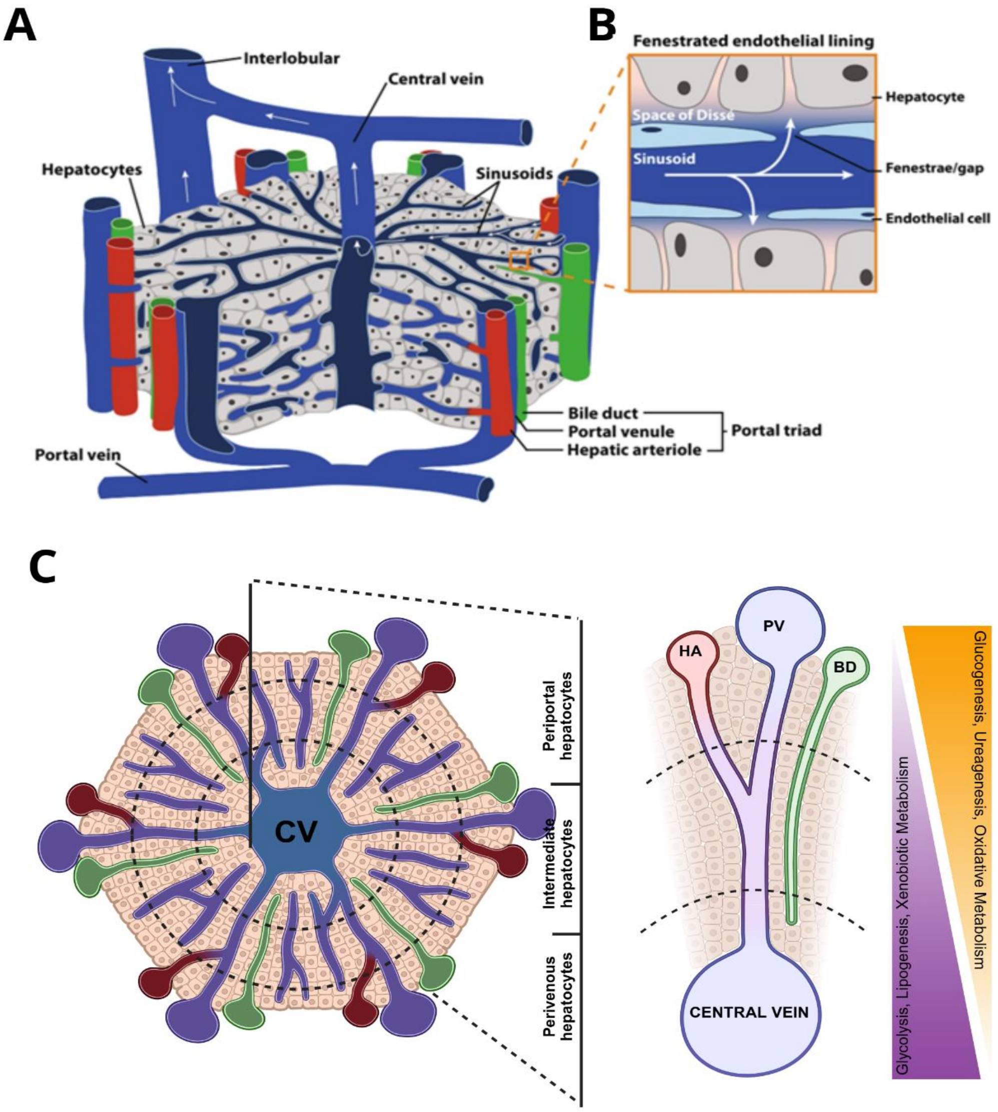


Figure 1.1: Structure of the hepatic lobule and metabolic zonation within the acinus.

(A) Schematic of the hepatic lobule illustrating vascular organisation and directional blood flow from the portal triad towards the central vein. (B) Enlarged representation of the sinusoidal endothelium highlighting fenestrations that permit diffusion of macromolecules into the Space of Disse surrounding hepatocytes. (C) Comparative view of the lobular (left) and acinar (right) organisation of the liver. In the lobular model, hepatocytes are arranged in concentric functional zones around the central vein (CV), corresponding to periportal, intermediate and pericentral regions. These zones arise from gradients in oxygen and nutrient availability and underpin hepatocyte metabolic specialisation. The acinar model emphasises perfusion-based zonation along the portal vein-central vein axis. BD, bile duct; CV, central vein; HA, hepatic arteriole; PV, portal venule. Image adapted from (18).

`

1.5 Inborn errors of metabolism

Inborn errors of metabolism (IEM) are a heterogenous group of disorders involving failure of the metabolic pathways involved in either breakdown or synthesis of carbohydrates, fatty acids, and proteins. Although any single condition is rare, taken together, IEM affect approximately 1 in 2500 live births (59). They result from mutations in single genes encoding for enzymes that catalyse the conversion of substrates into metabolites, with disease arising either from the accumulation of toxic substrates or the deficiency of essential metabolites. Most IEMs are inherited in an autosomal recessive manner, and management commonly involves dietary therapy, with liver transplantation required in selected cases. However, a lack of available donors, risks associated with the lifelong immunosuppression, and the potential for irreversible damage before transplantation becomes available remain major limitations. Among the best-studied IEM is phenylketonuria (PKU), which exemplifies how a defect in a single hepatic enzyme can lead to profound systemic disease.

1.6 Phenylketonuria

Phenylketonuria (PKU) is an autosomal recessive disorder of phenylalanine metabolism, characterised by a deficiency in phenylalanine hydroxylase (PAH), the enzyme responsible for metabolising phenylalanine to tyrosine. The resulting accumulation of phenylalanine in the blood and brain, termed hyperphenylalaninemia, is highly neurotoxic. Although infants with PKU are healthy at birth, untreated individuals rapidly develop severe intellectual disability, developmental delays, psychiatric and movement problems, microcephaly and epilepsy as well as light pigmentation of skin, eyes and hair and a characteristic musty odour (60).

`

The condition was first described in 1934 by Dr Ivan Asbjørn Følling, who investigated two Norwegian siblings with severe mental disability and peculiar odour (61). During routine urine testing for ketones, he observed an unusual reaction that was later identified as elevated phenylpyruvic acid. Similar findings were subsequently noted in other institutionalised individuals with intellectual disability. In 1953, Jervis demonstrated that PKU resulted from a metabolic block in phenylalanine metabolism, and later that same year Bickel et al. showed that dietary restriction of phenylalanine improved clinical outcomes, thereby establishing the foundation of treatment for PKU (62,63).

The pivotal introduction of population-based newborn screening soon followed. Dr Robert Guthrie, paediatrician and father of a child with PKU, pioneered the dried blood spot (DBS) assay to measure phenylalanine concentrations from a simple heel-prick blood sample shortly after birth (64). This innovation allowed for early diagnosis and timely intervention, which if adhered, prevented severe intellectual disability. The DBS assay has since been adapted to screen for a wide range of metabolic diseases at birth and is also used to monitor blood phenylalanine levels in PKU patients, maintaining concentrations within a healthy range (65). According to the Australian consensus guidelines for the management of PKU, blood phenylalanine should be maintained between 120-360 µmol/L in children and in women pre-conception or during pregnancy, and between 120-600 µmol/L in individuals over 12 years old (66). In contrast, untreated individuals with PKU can often exhibit blood phenylalanine concentrations exceeding 1200 µmol/L (67). Dietary protein tolerance is defined as the intake that maintains blood phenylalanine level with the prescribed target treatment range and is dependent on several factors including the degree of PAH impairment, growth rate, age and gender (68). On average, adult patients can tolerate approximately 4 to 6 g of protein (200 to 300 mg Phe)/day (67). As a result, high-protein foods such as meat, eggs, dairy wheat, rice and other grains must be strictly avoided, as well as artificial sweetener aspartame which is metabolised into phenylalanine. In addition to severe dietary restriction, patients require amino acid supplementation to meet nutritional needs of the other amino acids. These supplements are often unpalatable, must be consumed frequently, and impose a significant financial burden (69). Special phenylalanine-free versions of common foods, such as pasta and bread, are also available and allow individuals to have a more typical diet, but these too are costly and contribute further to the economic and practical challenges of lifelong dietary management (67). For these reasons, dietary compliance is difficult, and nutritional deficiencies are common (70).

Even among diet adherent patients, rates of attention deficit, specific learning disabilities, anxiety and depression remain higher than in the general population (71). Women with PKU face unique challenges during pregnancy, as elevated maternal phenylalanine is a potent teratogen that can cause microcephaly, congenital heart defects, growth restriction, and intellectual disability in the foetus if strict metabolic control is not maintained (72). Adverse outcomes have been reported even in women who were previously well controlled, underscoring the need for intensive monitoring and strict dietary management before conception and throughout gestation. Despite the success of newborn screening and dietary management, substantial lifelong morbidity and psychosocial impacts associated with PKU remain a significant global challenge across populations.

1.6.1 Epidemiology

PKU is among the most common IEMs, with an estimated global prevalence of 1 in every 24,000 live births and 450,000 people living with PKU globally (1). However, prevalence

`

varies markedly between populations and geographic regions, reflecting both genetic variation and the implementation of national newborn screening programs. In Australia, the incidence is around 1 in 10,000 live births, with approximately 30 new diagnoses made each year (70). According to Hillert et al., (2020)(1), the highest reported incidence occurs in Italy, where PKU affects 1 in 2,700, whereas countries such as Finland and Japan have extremely low rates, with fewer than 1 in 100,000 births affected. High incidences in countries such as Türkiye and parts of the Middle East are partly attributable to consanguinity and founder mutations, while East Asian populations more commonly present with milder forms of the condition rather than classical PKU (1). Although neonatal screening has enabled robust epidemiological data in most high-income countries, coverage remains incomplete in many parts of Africa, Asia, and the Caribbean, meaning true prevalence in these regions is still uncertain (73).

1.6.2 Phenylalanine metabolism

PKU arises from defects in phenylalanine metabolism. Phenylalanine is an essential amino acid required for protein synthesis and is acquired either through dietary intake of proteins along with endogenous recycling of amino acid stores. Under normal physiology, phenylalanine homeostasis is tightly regulated to balance incorporation into proteins with catabolic processes (74) (Figure 1.2).

The primary route of phenylalanine metabolism is through oxidisation to tyrosine by the enzyme phenylalanine hydrolase (PAH), which requires the cofactors tetrahydrobiopterin (BH4), molecular oxygen and ferrous iron. This reaction account for approximately 80% of metabolism and occurs primarily in the liver but also at low levels in the kidney (75), although neither organ shows any apparent pathology in PKU. Two alternative pathways of phenylalanine metabolism include the transamination to phenylpyruvate and decarboxylation to phenylethylamine; together these pathways represent less than 5% phenylalanine metabolism under normal conditions (74). It is estimated approximately10-20% of dietary phenylalanine is incorporated into proteins, however this proportion varies depending on physiological and nutritional factors (74,76). Tyrosine has several metabolic fates, including incorporation into proteins, specifically the production of neurotransmitters (dopamine, adrenaline and noradrenaline), synthesis of thyroxine in the thyroid gland, melanin in melanocytes and other derivatives, such as melatonin, and eventually complete catabolism into acetoacetate and fumarate for energy production (60).

When this primary pathway is impaired, as in PKU, flux through the alternative pathways increases but remains insufficient to compensate for the loss of PAH activity. As a result, phenylalanine accumulates systemically, and increased amounts are shunted towards the production of phenylpyruvate and phenylethylamine and rapidly excreted in the urine, specifically, phenylpyruvate and its secondary metabolites phenylacetate and phenyllactate, collectively known as phenylketones, give urine a characteristic musty odour from which the condition derives its name. Despite increase activity of these minor pathways, phenylalanine levels remain elevated, leading to neurotoxicity and establishing the biochemical basis of PKU.

`

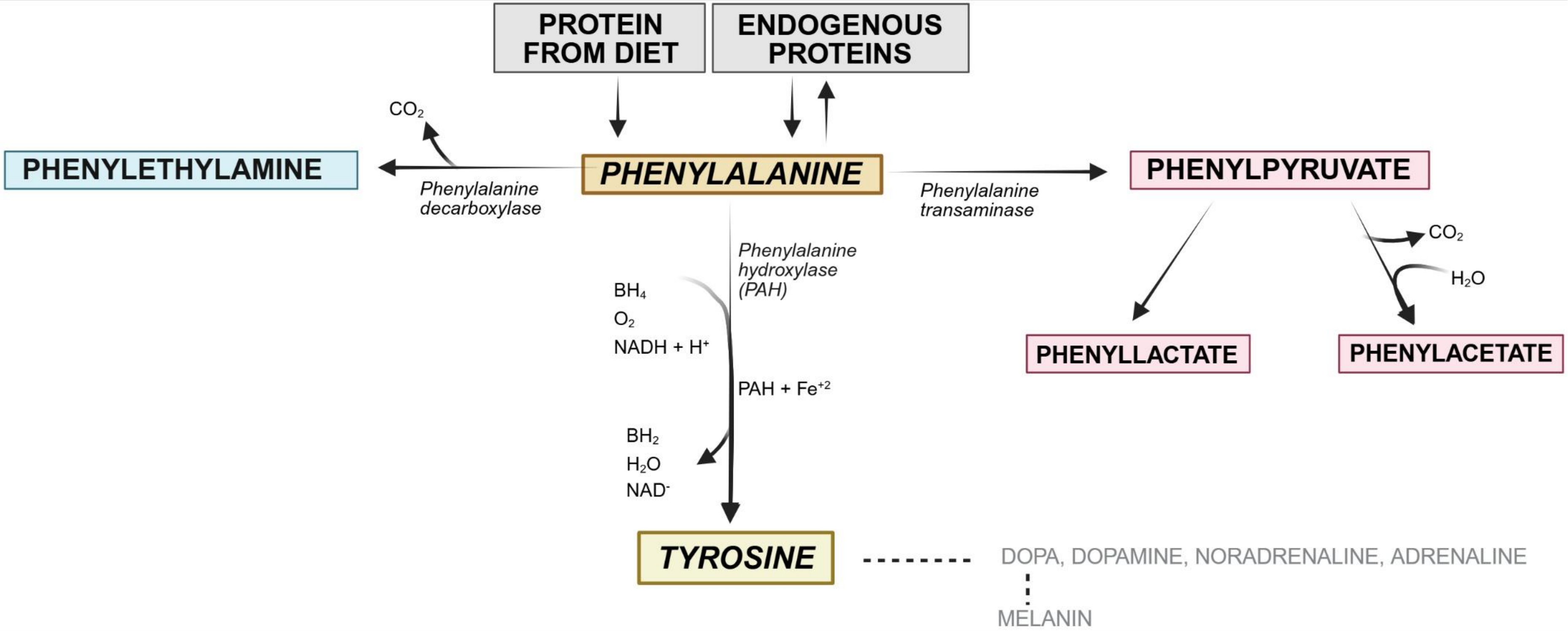


Figure 1.2: Metabolism of phenylalanine

Phenylalanine, obtained from dietary protein intake or endogenous proteolysis, is metabolised through several pathways. The primary pathway (~80% of normal metabolism) involves hydroxylation to tyrosine, catalysed by phenylalanine hydroxylase (PAH) in the presence of the essential cofactor tetrahydrobiopterin (BH_4), molecular oxygen, reduced nicotinamide adenine dinucleotide, and ferrous iron. Tyrosine subsequently serves as a precursor for the synthesis of neurotransmitters dopamine, noradrenaline and adrenaline, as well as melanin. Minor pathways (<5% of normal metabolism) include transamination to phenylpyruvate and its secondary metabolites phenyllactate and phenylacetate, metabolites collectively known as phenylketones (shown in red) as well as decarboxylation to phenylethylamine. Dietary phenylalanine is also incorporated into proteins (~10-20% normal metabolism). In individuals with phenylketonuria (PKU), deficient PAH activity prevents efficient conversion of phenylalanine to tyrosine, resulting in elevated phenylalanine in the blood and increased production of phenylketones which are rapidly excretion in urine – an observation from which the disease derives its name. These alternative pathways are insufficient to prevent hyperphenylalaninemia and its associated neurotoxicity. Modified from (74).

`

1.6.3 Neuropathology

Elevated phenylalanine concentrations exert their most profound effects the nervous system, where they disrupt brain development and function. However, the molecular basis of neuropathogenesis remains incompletely understood. Post-mortem studies of untreated individuals with PKU have shown reduced brain size, white matter abnormalities and impaired myelination (77). However, the most consistent biochemical alternations identified in PKU brains are (i) elevated phenylalanine (ii) decrease in other large neutral amino acids' tyrosine and tryptophan and (iii) the depletion of monoamine neurotransmitters dopamine and serotonin (78–80).

Several mechanisms have been proposed to explain PKU-associated neuropathology. High blood phenylalanine has been shown to compete with tyrosine and tryptophan for the same membrane channel (LAT-1) for transport across the blood-brain barrier (60). This leads to cerebral deficiencies of tyrosine and tryptophan. As these amino acids are precursors for dopamine, adrenaline, noradrenaline, acetylcholine and serotonin, their reduced availability impairs neurotransmitter synthesis having profound effects on the brain, especially the developing brain (81).

Other mechanisms have been identified including a direct inhibitory effect on tyrosine and tryptophan hydroxylase (82), dysregulation of cerebral metabolism (83), alterations to genome methylation and transcriptome (84) as well as the formation of phenylalanine amyloid-like plaques in the brain (85,86) and B-amyloid and tau in the CSF (80).

1.6.4 Genetic basis of PKU

PKU is an autosomal recessive condition caused predominantly by mutations in the gene encoding phenylalanine hydroxylase (*PAH*). A small subset of cases (<2%) are due to mutations in genes involved in the synthesis or recycling of the cofactor BH4 or in proteins that interact with PAH, leading to secondary PAH deficiency (60). More recently, aberrations in long non-coding RNA (lncRNAs) shown to facilitate PAH expression and/or function have been identified in both PKU animal models and human *in vitro* cultures but have yet to be reported in the clinic (87). The rest of this review will focus on PKU caused by mutations in *PAH*.

The *PAH* gene is located on chromosome 12 (12q22-q24.2), spans ~90kb, and contains 13 exons (60). Its coding sequence is 1,359 bp, encoding a 452-amino acid polypeptide monomer, which forms a tetrameric protein (Figure 1.3 A-C). The polypeptide sequence is organised into three distinct domains: an N-terminal regulatory domain, a central catalytic domain and a C-terminal tetramerisation domain.

Pathogenic variants are highly heterogenous, with 3369 disease causing mutations identified so far (88). Most patients are compound heterozygotes, leading to thousands of different distinct genotypes (Figure 1.3 D). It can be hard to predict properties of 2 different variant alleles, examples of both positive and negative inter-allelic complementation are known (73,89), making genotype-phenotype predictions difficult. Pathogenic variants are distributed across the entire gene (Figure 1.2 D). As of 2024, 43.4% of all pathogenic variants were intronic (77.0% of these near splice regions), 33.7% were missense variants, 10.0% were frameshifts, 7.7% synonymous substitutions, 2.9% nonsense variants and 1% were in-frame variants (88). Mutations in PAH are typically hypomorphic, leading to reduced or absent PAH

`

enzymatic activity. Baseline blood phenylalanine and dietary protein tolerance is directly related to the amount of residual function of PAH. The goal therefore of any therapeutic intervention is to clear excessive phenylalanine in the system, typically by increased functional PAH enzyme activity.

1.7 Existing therapies for PKU

Protein restriction combined with medical supplementation has remained the standard of care for PKU for more than 70 years. Efforts to improve treatment and quality of life have yielded two major pharmacological breakthroughs that enhance dietary tolerance to phenylalanine. The first is sapropterin dihydrochloride (marketed as Kuvan ®), a synthetic form of the BH_4 cofactor administered orally daily which was approved in 2007 (90). Sapropterin is typically effective in individuals with residual PAH activity (91), with approximately 20-60% of patients achieving >30% reduction in blood phenylalanine levels. In some cases, patients were able to discontinue dietary restriction altogether (90,92,93) Responsiveness largely depends on the underlying PAH mutations, with most responsive cases involving variants in the catalytic domain of PAH (89). Nevertheless, up to 80% of individuals may show little to no response to sapropterin, and the treatment’s considerable cost – estimated between 48,000 and 169,000 USD annually for a 68 kg patient (94) – further limits its accessibility.

`

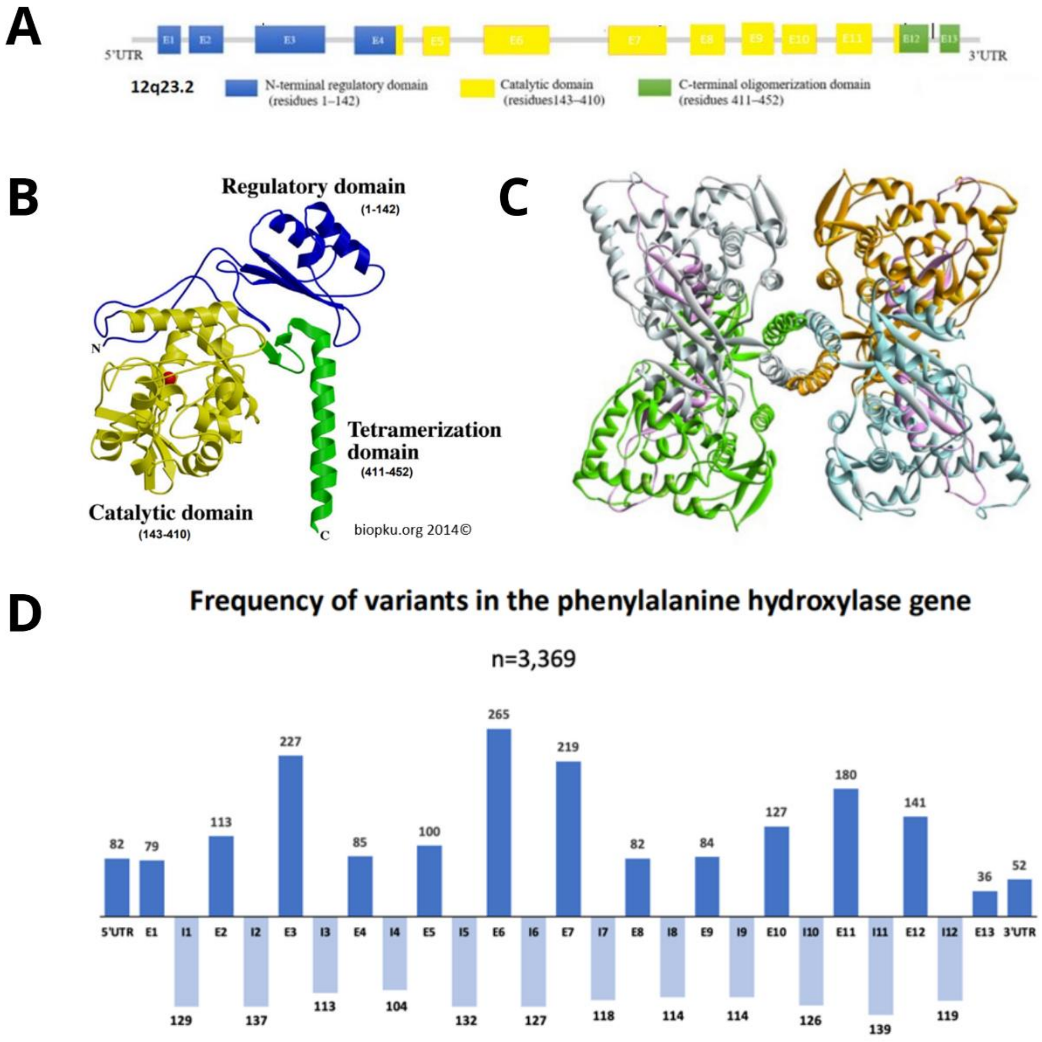


Figure 1.3: Genetic basis of PKU

(A) Schematic of the PAH gene (chromosome 12q23.2), showing the 13 exons and corresponding protein domains: N-terminal regulatory domain (residues 1–142), catalytic domain (residues 143–410), and C-terminal oligomerisation/tetramerisation domain (residues 411–452). (B) Cartoon representation of the PAH monomer, highlighting the regulatory, catalytic, and tetramerisation domains (95). (C) Tetrameric structure of full-length PAH, adapted from (96). (D) Distribution of reported variants in the *PAH* gene (n = 3,369), showing variant frequency per exon (light blue) and across untranslated regions and introns (dark blue)(97).

`

In 2018, the FDA approved pegvaliase (Palynziq ®), an enzyme substitution therapy consisting of a PEGylated recombinant phenylalanine ammonia lyase (PAL) derived from cyanobacteria. Pegvaliase metabolises circulating phenylalanine into trans-cinnamic acid and ammonia (98), effectively lowering blood phenylalanine to levels that allow some patients (up to 70%) to relax or even discontinue dietary restrictions (99,100). However, the therapy's daily subcutaneous injections and is associated with immune-mediated hypersensitivity reactions, including anaphylaxis in extreme cases, though the frequency and severity of these adverse events tend to decrease with continued use (101). Furthermore, the annual cost 192,000 USD, again is a substantial barrier to widespread accessibility (102). Together, sapropterin and pegvaliase have significantly improved outcomes and quality of life for subsets of PKU patients. Nonetheless, their limitations highlight the need for long-term, universally applicable therapies.

1.7 Models of PKU

Many models of PKU have been developed to better understand disease pathogenesis and to evaluate potential therapies beyond what can be learned from epidemiological studies. *In vitro* models provide controlled conditions to investigate fundamental biology in cell-specific manner but lack the complexity of whole organisms (103). By contrast, animal models provide a more clinically relevant model, capturing physiological interactions and enabling preclinical testing, but they are limited by species-specific differences and often fail to recapitulate the full spectrum of human pathology. Human tissue and humanised models offer opportunities to study therapies in a more physiologically relevant context but are constrained by limited access to patient material. A balanced approach that integrates *in vitro*, *in vivo* animal and humanised models is essential for a comprehensive understanding of gene therapy potential (104).

In addition, *in silico* tools based on curated data are valuable for predicting molecular and therapeutic outcomes, though these will not be covered in this chapter. Examples include tools for rapid characterisation of new PAH variants, AlphaFold for modelling structural consequences of individual mutations on PAH tetramers, algorithms predicting BH_4 responsiveness across patient genotypes and models assessing patient responsiveness to PAL-expressing probiotics (88,89,105,106).

1.7.1 *In vitro* models of PKU

In vitro models of PKU provide valuable systems to dissect the molecular consequences of specific PAH mutations. Several studies demonstrated that different PAH variants affect tetramer stability and biophysical properties in diverse ways (107–111). These models have also been instrumental for genotype-phenotype correlation studies, helping to link specific alleles to residual enzyme activity and clinical severity (112,113). More recently, advanced *in vitro* systems, including patient-derived brain organoids from induced pluripotent stem cells (iPSCs) that capture aspects of PKU neuropathology in a three-dimensional context (114). Liver organoids derived from iPSCs have already been generated for other metabolic conditions (115) and could be applied in the case of PKU.

1.7.2 Animal models of PKU

To explore the potential of novel therapies, an appropriate animal model that accurately recapitulates human disease is essential. Table 1 provides a summary of animal models of

`

PKU. The most widely used model by far is the *Pah*$^{enu2/enu2}$ mouse*,* generated by random chemical mutagenesis, which introduced a missense mutation in exon 7 (p.Phe263Ser, F263S). This substitution disrupts the catalytic domain of PAH, abolishing enzymatic activity while maintaining normal protein stability (116). Homozygous *Pah*$^{enu2/enu2}$ mice faithfully recapitulate the symptoms of untreated PKU in humans, including blood phenylalanine levels exceeding 1500 µM/L, hypopigmentation, and neurocognitive and behavioural impairments. Less well-documented phenotypes include osteopenia, altered serotonin and dopamine metabolites in the brain and increased rates of anxiety and depression-like behaviours (60). Homozygous females display maternal PKU, with significantly increased rates of spontaneous abortions and smaller litter sizes (117). This model has therefore become the cornerstone of preclinical testing and has been used extensively in therapeutic development for over 30 years.

One caveat of the *Pah*^{enu2} model is the expression of residual mutant PAH protein, which can interfere with any supplied PAH protein by forming hybrid wildtype and mutant PAH tetramers (118). This interaction may interfere with therapeutic intervention and/or the interpretation of preclinical studies. However, as most PKU patients' express mutant PAH protein, the feature retains significant clinical relevance by most accurately representing challenges facing treating PKU. Effective therapies must therefore compete with or complement existing alleles rather than function in their absence. This is challenge is compounded by the fact that PKU is caused by a highly heterogenous spectrum of mutations with no predominant allele across populations (1). Most patients are compound heterozygotes, carrying two different pathogenic variants that may impair PAH activity through different mechanisms (113).

`

To simplify preclinical testing by eliminating the confounding effects of residual protein, two groups have independently generated *Pah* knockout mice completing devoid of PAH protein, both of which recapitulate classic PKU phenotypes (119,120). Other models have been generated to answer specific questions regarding PKU pathogenesis. Aubi et al., (2021)(86) argue that PAH protein instability is the primary pathological mechanism of PKU and as such developed a knock-in mouse carrying a common missense mutation in humans, R261Q. These mice only exhibited elevated blood phenylalanine when fed a high-protein diet and did not develop motor disturbances. They did however form PAH amyloid-like aggregates in the liver, suggesting a possible toxic gain-of-function mechanism in addition to loss of enzymatic activity. Martinez-Pizarro and colleagues (2024) (121) introduced human intron 10 containing the c.1066-11G>A splice variant, producing the first splicing-defect PKU model. These mice exhibited absent PAH protein and no enzymatic activity despite normal mRNA levels, alongside classical PKY manifestations including elevated blood and brain phenylalanine. Zhang et al., (2025)(122) described another humanised knock-in model carrying an intronic mutation. These mice also developed hyperphenylalaninemia and demonstrated impaired spatial learning and memory, although molecular characterisation of PAH protein or enzyme activity were not reported.

Finally, large-animal models have been generated to improve clinical translatability. Porcine PKU models have been established independently by two groups and offer the advantage of modelling complex neurocognitive phenotypes and provide a platform for scaling therapeutic interventions (123,124).

1.7.3 Humanised models for gene therapy

Species differences can limit the utility of animal models in clinical translation, not only due

`

to obvious physiological differences but also because of more subtle molecular discrepancies. For example, AAV vectors used to deliver genetic cargo across the blood brain barrier use receptors which can differ between species, and even between different inbred mouse strains (125). It is therefore crucial to, where possible, to employ models that closely mimic human biology. Primary liver cells taken from cadavers, medical biopsies or explanted donor organs are currently the gold standard of experimental models but are difficult to work with due to their propensity to rapidly de-differentiate and loss of mature hepatocytes function within days, as well as the difficulty in obtaining sufficient material. Access to hepatocytes from PKU patients is particularly limited as liver transplantation is not part of the standard of care. An alternative approach involves the generation of iPSC-derived hepatocyte-like cells (iHeps) or human liver organoids. Although significant advances have been made in these fields, these models lack the full genetic profile of mature hepatocytes and do not completely replicate the complex, heterocellular, three-dimension architecture of the human liver (126,127).

One way to partially address these limitations is through xenotransplantation of human hepatocytes into animal hosts, such as the widely used double immune-deficient FRG (Fah-/-, Rag2-/-, Il2rg-/-) mouse model (128). In this model, the absence of fumarylacetoacetate hydrolase (FAH) results in a lethal defect in tyrosine metabolism that causes progressive liver failure. Mice are maintained on the protective drug 2-(2-nitro-4-trifluoromethylbenzoyl)-1,3-cyclohexanedione (NTBC), which blocks tyrosine catabolism upstream of FAH and prevents hepatotoxicity. Following transplantation of human hepatocytes, withdrawal of NTBC creates a selective pressure whereby murine hepatocytes undergo apoptosis while human hepatocytes to proliferate. Cycling on and off NTBC can results in up to ~90% repopulation of the murine liver with human cells. While patient-derived hepatocytes can be transplanted in cases of liver disease where transplantation is clinically indicated, this is not applicable to PKU. Looking forward, it may become possible to engineer hepatocytes carrying specific PAH mutations from cell lines, differentiate them, and engraft them into these humanised mouse models.

Recent advances in *ex vivo* liver perfusion technology have also enabled the maintenance of explanted human livers on circuit systems for extended periods. Cabanes-Creus and colleagues (2024) (129) utilised human livers deemed unsuitable for transplantation to evaluate AAV vector performance in the presence or absence of neutralising antibodies for several day post-infusion. This approach likely represents one of the most biologically and clinically predictive preclinical model, though it remains highly sporadic and limited in availability. These of course do not provide an entire systematic view, with PKU pathology manifesting in the brain. To explore the neuropathological consequences of elevated phenylalanine, hippocampal slice cultures have been used as an *ex vivo* model, for instance by culture tissue in the presence of PKU patient cerebrospinal fluid (130). For now, testing in animal models such as mice contribute to provide the most accessible and widely use platform for preclinical PKU research.

Table 1.1: Animal models of PKU help elucidate pathogenesis and allow therapeutic testing

Animal	Name	Mutation	PAH Protein	Notable presentation	Ref
Mouse - BTBR	Pah enu1	c.T364>C missense mutation in exon 3. Regulatory domain	Unstable monomer. Reduced expression of functional protein, 5% activity that of wildtype	Mildly elevated blood Phe with protein challenge	(116)
Mouse - BTBR	Pah enu2	c.T835>C missense mutation in exon 7. Catalytic domain	Protein expressed but not functional		(131)
Mouse - C57BL6/J	Pah enu2	c.T835>C missense mutation in exon 7. Catalytic domain	Protein expressed but not functional	Pah amyloid like aggregates in the brain	(131)
Mouse - BTBR	Pah enu3	c.1199 + 2T > G frameshift mutation resulting in premature stop codon. Tetramerization domain	No protein expressed		(132)
Mouse - C57BL6/N	Pah exon 1 deletion	Exon 1 deletion	No protein expressed	Prevents mutant protein interfering with any therapeutic protein introduced. Model provides easier read out for gene therapy interventions	(119)
Mouse - C57BL6/J	Pah R261Q	c.782G>A missense mutation. Common mutation in patients		PAH amyloid-like aggregates in liver. Used to explore toxic gain of function component of PKU by exploring PAH aggregates	(86)
Mouse - C57BL6/J	Humanised P281L	Replaced murine exon 7 with human exon 7 containing missense mutation c.842C>T		Humanised model of PKU	(133)
Mouse - C57BL6/J	Humanised c.1066-11G>A knock in	PAH splice variant using humanised intron 10 sequence to recapitulate splicing defect	No PAH protein produced. No PAH activity. Normal mRNA PAH levels	One of the most common PAH variants seen in patients. No patho-mechanism for this mutant protein is known	(121)
Mouse - C57BL6/J	Humanised c.1199+502 A>T knock in	Humanised mouse model of the most frequently reported deep intronic variant occurring in intron 11.	Nor reported	Humanised model of PKU	(122)
Mouse – C57BL6/N	Knockout Pair lncRNA	Pair exon 1 deletion	No changes to PAH protein levels	Alternative causes of PKU	(87)
Mouse – C57BL6/J	PAH knockout	Premature stop codon in codon 7	No protein expressed	Prevents mutant protein interfering with any therapeutic protein introduced. Model provides easier read out for gene therapy interventions	(120)
Minipig - Yucatan	PAH exon 6 deletion		No protein expressed	Large animal model of PKU. Closely resembles neuropathology observed in PKU human brain	(134)
Minipig - Ossabaw Pig - Yorkshire	Humanised p.R408W allele	Exon 8 affecting catalytic domain. Common and severe mutation in humans	Protein expressed but not functional	Large animal model of PKU	(124)

`

1.8 Novel Therapeutics for PKU

A range of approaches have been explored for the treatment of PKU, with current and recent clinical trials summarised in Table 1.2.

1.8.1 Novel drug approaches

Researchers are continuing to build on existing pharmacological therapeutic approaches to treat PKU. One approach is the oral delivery of PAL via probiotic *E.coli* strain, designed to metabolise phenylalanine within the enterohepatic circulation (135). A phase I/II trial in healthy volunteers and individuals with PKU demonstrated that while the probiotic was functional and well tolerated, administered doses failed to lower circulating phenylalanine levels (136). Dose-escalation studies are ongoing (NCT05764239). In parallel, the Children's Hospital of Fudan University, in collaboration with CommBio Therapeutics, has listed an upcoming phase I trial testing engineered probiotics in children with PKU (NCT05948020), although no supporting publications are currently available.

Another avenue of drug research is oral sepiapterin which is a natural precursor of BH4 that is being developed as an alternative to sapropterin therapy. Sepiapterin has been shown to increase plasma BH4 concentrations more effectively than sapropterin and to double BH4 levels in the cerebrospinal fluid (CSF). A phase III long-term efficacy trial is ongoing (NCT05166161) (137,138).

Finally, small-molecular inhibitors of the neutral amino acid transporter SLC6A19, responsible for renal reuptake of phenylalanine, are being explored to promote urine excretion. In Pah^{enu2} mice, treatment with an SLC6A19 inhibitor (250mg/kg) reduced plasma phenylalanine levels by approximately 50% within three hours, accompanied by a marked

`

increase in urinary excretion. However, the inhibitor lacks specificity, also enhancing the excretion of other amino acids, which may necessitate dietary supplementation and close monitoring (139). A phase I clinical trial is currently recruiting (NCT05781399). Although these pharmacological approaches show promise in alleviating hyperphenylalaninemia, both enzyme substitution and cofactor therapy require ongoing administration and do not provide a long term, definitive solution for PKU.

1.8.2 Cell therapy to treat PKU.

Another strategy under investigation for treating PKU is hepatocyte transplantation. Orthotopic liver transplants could cure PKU however given the disease is non-fatal and there is a limited availability of donor grafts restricts its use as a therapeutic option (140). Consequently, hepatocyte transplantation followed by liver repopulation has emerged as a potential cell-based therapy for PKU (141).

The successful implementation of allogenic hepatocyte transplantation faces several challenges, including the requirement of immunosuppression and the efficient initial engraftment and subsequent expansion of transplanted cells. Even with optimised protocols, more than 70% of infused cells are typically cleared with the first 24 hours post-infusion by host immune system (142). For sustained engraftment, transplanted hepatocytes must acquire a selective advantage. As the liver is structurally and functionally normal in PKU, there is not inherent advantage for donor cells. Selective pressure can be artificially induced through preconditioning regimens, such as preparative hepatic irradiation, which was used in the only hepatocyte transplantation clinical trial for PKU to date (NCT01465100). This trial enrolled a single participant and was terminated prematurely due to lack of funding, with no published results.

An alternative approach was explored in a preclinical study by Vonada et al., (2024)(143), who engineered wildtype hepatocytes *ex vivo* by knocking out *Cypor* using SpCas9 prior to transplantation into *Pah*$^{\mathrm{enu2/enu2}}$ mouse. *Cypor* is required for acetaminophen (APAP) metabolism into toxic metabolites, and thus knockout hepatocytes are resistant to APAP-induced liver injury. Following administration of APAP to create a selective pressure, donor cells expanded to ~14% of the liver mass, resulting in a sustained reduction of blood phenylalanine to physiological levels. This was stable after the withdrawal of APAP selection and saw no impairment of overall hepatic metabolism. Despite these promising results, immunosuppressive therapy remains essential to prevent graft rejection, bringing inherent risks that continue to limit the clinical applicability of hepatocyte transplantation for PKU (140).

1.8.3 Gene therapy to treat PKU

There has been significant interest in developing gene therapies for PKU. Both viral and LNP vectors have been explored to deliver functional PAH DNA or mRNA to hepatocytes or to elicit permanent changes to genome to recover wildtype PAH expression.

1.8.3.1 LNP-mRNA approaches in development to treat PKU

Successful preclinical studies demonstrated correction of hyperphenylalaninemia and restored liver PAH activity in *Pah*$^{\mathrm{enu2/enu2}}$ mice treated with either bacterial PAL mRNA, murine PAH mRNA, human PAH mRNA delivered in different proprietary LNPs (144–147). However, these effects were short-lived, with levels returning to pre-treatment levels within 48 hours. Thus, repeating dosing is required for therapeutic effect, raising concerns about immunogenicity, cost and practicality. Chemically modified and circular RNAs are being

`

developed to improve stability, prolong *in vivo* persistence, and extend the therapeutic window (148). Moderna had planned to test an LNP-PAH mRNA therapy in a combined phase I/II clinical trial (NCT06147856) based on preclinical successes (144), but this study has since been withdrawn.

1.8.3.2 AAV gene addition approaches in development to treat PKU

Delivery of PAH therapeutic DNA constructs to hepatocytes, otherwise known as gene addition, has been achieved in many ways over the past three decades since the *PAH* gene was cloned. Approaches have used different viral vectors (AdV, lentivirus, AAVs), AAV capsids (AAV1, AAV2, AAV8, AAV-Anc80, AAVHSC15), promoter sequences, and *PAH* transgene cargos (murine *Pah*, human *PAH*, codon optimised *PAH*) (124,149–157). These approaches have achieved strong preclinical correction in the Pah^{enu2} mouse model (141), leading to multiple clinical programs, though with limited clinical success to date. Homology Medicine tested a proprietary AAV vector (AAVHSC15) to deliver a codon-optimised human *PAH* cDNA, which normalised blood phenylalanine in Pah$^{enu2/enu2}$ mice (150). This progressed to the *pheNIX* trial (NCT03952156) but was placed on hold by the FDA in 2022 due to elevated liver function tests and later discontinued (158). Limited interim results from part of the dose-escalation trial were reported: of six treated patients, one individual in the mid-dose cohort experienced a reduction in blood phenylalanine from 1010 μmol/L at baseline to 42 μmol/L after 48 weeks, and one individual in the high-dose cohort decreased from 1060 to 303 μmol/L by 13 weeks. No serious adverse events were reported, and elevated liver enzymes were managed by corticosteroids.

BioMarin's *Phearless* trial (NCT04480567) employed an AAV5 to deliver PAH cDNA. The program was paused in 2021 following concerning preclinical observations in a PAH-null,

`

immunodeficient mouse model (NRG/PAH$^{-/-}$), which was selected to enable long-term persistence of AAV genomes in the absence of adaptive immune clearance. In this model, 6 of 7 mice treated at the highest dose (2×10^{14} vg/kg) developed adenomas or hepatocellular carcinoma at necropsy, with evidence of AAV integration into the genome (159). To date, only lower vector doses (2×10^{13} vg/kg and 6×10^{13}vg/kg) have been administered in human participants. Following these findings, the FDA requested additional preclinical safety studies, and the clinical trial remains ongoing with an anticipated completion in 2027.

More recently, NGGT Inc. and Sanofi have announced new programs using AAV8 and a proprietary AAV capsid respectively in adult cohorts. While these efforts highlight sustained interest, a clear clinical benefit in humans has not yet been demonstrated. Several factors may contribute to the limited success to date, such as suboptimal hepatocyte transduction efficiency of certain AAV capsid in humans *in vivo* and insufficient promoter activity. A further critical consideration is the proliferative status of the liver, which poses a major barrier to the durability of gene addition strategies in PKU.

1.8.3.2 Durability of AAV gene addition in the liver

A major limitation of AAV-mediated gene addition is that transgene expression relies on episomal vector genomes, which are progressively lost with each cell division. In non-dividing or slowly dividing tissues, such as muscle or neurons, episomes can persist for many years (160,161). In contrast, hepatocyte proliferation is particularly pronounced during early development, when the liver undergoes rapid growth, approximately doubling in relative size by the approximate age of 6 months, 2 years and 9 years (162). Consistent with this, rapid loss of AAV episomes in neonatal mice has been demonstrated using GFP reporter constructs (163).

Although this rapid proliferative phase subsides as hepatocytes become more quiescent in adulthood, hepatocytes continue to renew throughout life with an average cell age of less than three years in adult humans (164). Importantly, hepatocyte turnover is not uniform across the liver lobule but is zonated, with higher rates of proliferation reported in periportal region. This ongoing cellular turnover even in adulthood contributes to the gradual dilution and loss of episomal AAV genomes over time, which may ultimately result in waning therapeutic efficacy.

Despite these limitations, multiple preclinical and clinical studies have shown the utility of AAV gene addition in the liver, with sustained transgene expression observed for years in adult hepatocytes. Notably, durable expression has been reported in haemophilia B dog models for over eight years of follow up (165), as well as in human patients following systemic AAV8-FIX infusion (3-6 years follow up) (37,166). These findings suggest that, in some contexts, episomal persistence in adult hepatocytes may be sufficient for long-term therapeutic benefit, particularly in genes encoding secreted proteins or for pathways that are not strongly affected by hepatic zonation.

By contrast, metabolic functions that are preferentially localised to periportal hepatocyte, such as the urea cycle, may be more prone to loss of efficacy over time, even in adult cohorts. Moreover, re-dosing AAV vectors is currently not feasible due to the development of humoral immune responses against the viral capsid, and the prevalence of naturally acquired pre-existing immunity increases with age (33). Collectively, these considerations have motivated the exploration of gene editing approaches that are stably inherited during hepatocyte division and therefore offer the potential for more durable therapeutic correction.

`

1.8.3.3 Gene editing to treat PKU

Gene editing seeks to introduce permanent modifications to the endogenous genome. This ensures therapeutic changes are stably inherited during hepatocyte proliferation, addressing the durability limitations of AAV and LNP gene addition approaches. Gene editing can be achieved by programmable nucleases such as CRISPR/Cas9, which introduces double-stranded DNA breaks (DSBs) that are repaired by the cell's own DNA repair machinery.

1.8.3.3.1 CRISPR/Cas9 gene editing

Clustered regularly interspaced short palindromic repeats (CRISPR) are part of a prokaryotic adaptive immune system, in which fragments of invading viral genomes are incorporated into host CRISPR arrays and subsequently used to guide sequence-specific cleavage of foreign DNA (167–169). This discovery led to the adaptation of the *Streptococcus pyogenes* CRISPR-Cas9 system (SpCas9) into a programmable genome editing tool capable of functioning in eukaryotic cells (170–172).

Compared with earlier gene editing technologies such as zinc-nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs), CRISPR-Cas9 offers a simpler and more modular targeting strategy. Specificity is dictated by a short single-guide RNA (sgRNA) of ~20 nucleotides which hybridises to the target DNA sequence adjacent to a protospacer adjacent motif (PAM). The sgRNA directs the Cas9 endonuclease to produce a double-stranded break, typically 3 bp upstream of the PAM (170).

Since the initial development of the SpCas9 system, a range of orthologues have been identified and engineered to expand the CRISPR toolkit. A notable example is *Staphylococcus aureus* Cas9 (SaCas9) which is widely used in gene therapy applications due to the smaller coding sequence (173), allowing for the entire sequence to be packaged into a single AAV vector (174). Many orthologues have been identified. Other examples include Cas9's with distinct PAM requirements, allowing access to genomic site not targetable by SpCas9, such as St1Cas9 (175). In addition, nuclease inactive Cas9 (dCas9) binds DNA without cleaving it and has been repurposed for applications such as transcriptional regulation and epigenetic modulation (176,177).

1.8.3.3.2 DSB repair mechanisms are exploited for gene editing

Cas9-mediated DNA cleavage can be exploited to modify the genome by activating the cell's endogenous DNA repair machinery to introduce changes. These can take the form of This enables a range of genome engineering outcomes, including mutation correction, disruption of pathogenic alleles or targeted insertion of therapeutic sequences. Following Cas9-induced double stranded breaks (DSBs), cells primarily employ two repair pathways, non-homologous end joining (NHEJ) and homology directed repair (HDR). Both pathways can be harnessed to achieve specific gene editing outcomes, albeit with distinct constraints and efficiencies (Figure 1.4) (178).

NHEJ is the predominant DSB repair mechanism and is active throughout the cell cycle (Figure 1.4) (179,180). This pathway involves the direct ligation of broken DNA ends and can repair DSBs either faithfully or in an error-prone manner, frequently resulting in small insertions and deletions of DNA bases (indels) at the break site (179). This introduction of indels can be exploited to disrupt gene function by introducing frameshifts, exon skipping, or other premature stop codons that ablate gene functionality. As a result, NHEJ-mediated gene disruption has been widely applied in both experimental and clinical contexts for the treatment of dominant-negative or gain of function disorders (14,53). A prominent clinical example is Casgevy, an *ex vivo* CRISPR/Cas9 gene therapy exploiting NHEJ to disrupt erythroid-specific enhancer of BCL11A, thereby reactivating foetal haemoglobin expression for the treatment of sickle cell disease and β-thalassemia (14).

Beyond gene disruption, NHEJ can also be adapted for targeted integration of exogenous DNA through approaches such as homology-independent targeted integration (HITI) (Figure 1.4) (181). In HITI, both the genomic target site and the donor DNA are cleaved by Cas9 using the same guide RNA, generating compatible DNA ends that can be ligated via NHEJ. Importantly, the donor construct is designed such that correction orientation of integration disrupts the Cas9 cleavage, whereas reverse integration preserves it. If unchanged, this leads to repeated cleavage and the favouring of the correct insertion orientation. This strategy enables directional and site-specific integration without the need for homology arms and is therefore effective in non-dividing or slowing dividing cells, such as hepatocytes. These properties make HITI particularly attractive for *in vivo* gene therapy applications and has been successfully applied to correct autosomal dominant retinitis pigmentosa in mouse and pig retina (171,172), to integrate transgenes into the albumin safe harbour *in vivo* (173) and in the *ex vivo* manufacture of CAR-T cells (174).

Another DSB repair mechanism, HDR, represents a distinct repair mechanism capable of precise correction but requires a DNA template with homology to sequences flanking the break to guide accurate repair (Figure 1.4). This template may be the sister chromatid or a supplied donor sequence containing homology arms. Unlike NHEJ, HDR is restricted to dividing cells and is most active during S and G2 phases of the cell cycle, when sister chromatid is available as a template (178). In the context of the liver, this requirement significantly limits the proportion of hepatocytes that can be edited at any given time. While

HDR enables precise introduction of single nucleotide variants or insertions of therapeutic cassettes, its clinical utility has been constrained by competition with NHEJ and its dependence of cell proliferation (182).

1.8.3.4 Next generation genome editors

Recent iterations of CRISPR/Cas9 technology have sought to overcome the limitations of DSB-dependent editing through the development of base editors (183–185) and prime editors (186). These tools allow precise genomic changes without the need for double stranded DNA cleavage or exogenous donor DNA. Base editors allow the direct conversion of single nucleotides while prime editors integrate small sequences (<200 bp efficiently) (187,188). Although highly effective for use in disorders caused by common mutations, PKU present as greater challenge, with over 1,000 causative mutations identified in the *PAH* gene; and most patients are compound heterozygotes. In this context, base or prime editing would require bespoke reagents for each patient genotype. Given our current medical system where therapeutics are designed and tested for application across broad patient groups this extent of precision medicine is not practical, however, it may be attainable in the future.

1.8.3.5 Gene editing target sites

An important consideration for gene editing is where exactly to target. In monogenic conditions, it seems straightforward to modify the causative gene responsible for the disease. As described, this could include disrupting gene function, correct a single base pair mutation or introduce a several base pair therapeutic sequences. This has the distinct benefit of having the gene product under the control of endogenous promoter and enhancer sequences, with the possible risk of disrupting hypomorphic alleles depending on editing strategy employed.

`

Alternative strategies include integrating a therapeutic cassette into a well characterised genomic ‘safe harbour’ such as AAVS1 or CCR5 (170) or driving expression from a highly expressed loci such as albumin. The latter approach has already been explored for Haemophilia A and B as well as methylmalonic acidaemia (189,190). In these cases, the therapeutic cassette was placed under the control of the endogenous albumin promoter, which is both strong and liver specific. Targeting albumin also enables expression of a transgene that can be used for secreted and non-secreted proteins, as the first exon encodes a signal peptide that directs protein export. While partial loss of albumin is generally well tolerated, it nonetheless introduces a safety consideration. Moreover, such strategies may risk ectopic express or suboptimal transcriptional regulation of the therapeutic insert.

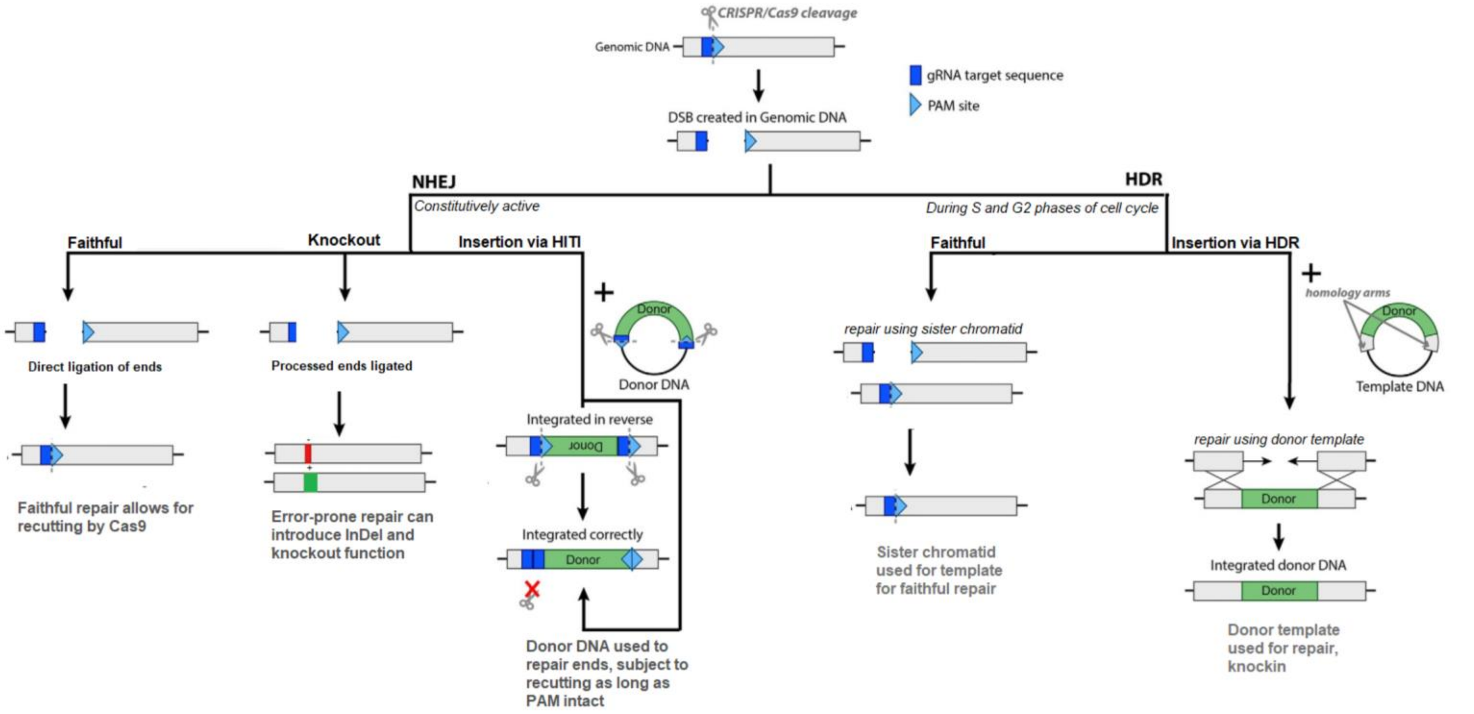


Figure Error! No text of specified style in document.**4: Overview of CRISPR-Cas9 genome editing outcomes following double strand break formation (DSB)**

CRISPR/Cas9 generates a blunt-end DBS at a specific genomic site (dark blue) adjacent to a PAM site (light blue), which can be exploited to edit DNA. DSB can be repaired in two main ways: Non-homologous end joining (NHEJ) and Homology-directed repair (HDR), with the pathway choice influenced by the cell cycle stage. NHEJ rapidly resolves DSBs by directly ligating DNA ends, which can restore the original sequence by faithful repair but more often introduces small insertions or deletions (indels) that disrupt gene function. NHEJ can also incorporate a donor sequence via homology-independent targeted integration (HITI). HITI donor sequences are flanked by CRISPR/Cas9 cleavage sites and may integrate in either orientation; sequences integrated in reverse can reconstitute PAM sites and be recut by Cas9, whereas correctly integrated sequences are resistant to further cleavage. Homology-directed repair (HDR) uses sister chromatid as a homologous template for precise repair and is therefore restricted to the S and G2 phases of the cell cycle. This pathway can be exploited for genome editing by providing a donor sequence containing homology arms which can be incorporated for DSB repair. Image modified from (18) with permission.

`

1.8.3.6 Clinical and preclinical gene editing approaches to treat PKU

To date, only one gene editing strategy for PKU has progressed to clinical evaluation (Table 1.2). Homology Medicines' developed HMN-103, a universal AAV-based gene editing approach designed to integrate a full-length *PAH* cDNA into intron 1 of the endogenous *PAH* locus via nuclease-free homologous recombination. The therapeutic cassette was delivered using the engineered AAVHSC15 capsid and relied on HDR to achieve targeted integration downstream of the native *PAH* promoter.

Preclinical evaluation in FRG human-liver chimeric mice demonstrated on-target integration in ~6% of alleles, without evidence of ITR co-integration or de novo mutations, although achieving this required a relatively high vector dose of 1×10^{14} vg/kg (191). Based on these results, a first-in-human clinical trial was initiated (NCT05222179). Three participants were enrolled: one individual exhibited a dramatic 98% reduction in blood phenylalanine from baseline to within a healthy range (<360 µmol/L), which was sustained even after dietary protein liberalisation at 31-week post-dose; a second participant showed a more modest reduction of ~49%; while the third experienced an unexpected increase in phenylalanine following vector administration. Despite these mixed but encouraging early outcomes, the programme was discontinued in 2023 due to financial constraints and the anticipated lengthy developmental timeline. Gene editing studies to treat PKU are now expanding, especially those exploring the application of base and prime editors. Preclinical editing studies are summarised in Table 1.3.

`

Table 1.2: Recent and ongoing clinical trials for PKU.

Therapy	Trial #	Delivery	Cohort	Mechanism of action	Stage	Company	#
Enzyme substitution	NCT05764239	Oral	Adults (≥18years)	SYNB1934: Oral delivery of PAL via probiotic *E. coli*	Phase III. **Terminated** June 2024	Synlogic	(135,136)
Enzyme substitution	NCT05948020	Oral	Children (3-17 years)	CBT102-A: Engineered probiotics that metabolize Phe by expressing related exogenous protein in Phe metabolic pathway (no further information given)	Phase I Recruiting	CommBio Therapeutics	
Cofactor therapy	NCT05166161	Oral	All ages	PTC923: sepiapterin, a precursor of tetrahydrobiopterin, to improve enzyme activity in mutant PAH	Phase III	PTC Therapeutics	(137,138)
Small molecule inhibitor	NCT05781399	Oral	Adults (18-65)	JNT-517: Small molecular inhibitor of neutral amino acid transporter, SLC6A19, responsible for Phe reuptake in the kidneys. Excess Phe excreted in urine	Phase I	Janana Therapeutics	(139)
Cell therapy	NCT01465100	Surgical	Adults	Hepatic irradiation followed by rounds of hepatocyte transplantation.	Phase I. **Discontinued** June 2023	Ira Fox	
LNP mRNA addition	NCT06147856	IV	Adults (18-70)	mRNA-3210: LNP containing PAH mRNA	Phase I/II. **Withdrawn** October 2024	Moderna TX, Inc.	(144)
AAV Gene addition "pheNIX"	NCT03952156; NCT04348708	IV	Adults (18-55)	HMN-102: AAVHSC15 containing PAH cDNA	Phase I/II **Discontinued** August 2023	Homology Medicine Inc	(150)
Gene addition "Phearless"	NCT04480567	IV	Adults and children (15+)	BMN307: AAV5 containing PAH cDNA	Phase I/II. Active, not recruiting (Hold)	BioMarin Pharmaceutical	
Gene Addition "PHEdom"	NCT06332807 (US); NCT06061614(China)	IV	Adults (18-55)	NGGT002: AAV8 carrying cDNA	Phase I/II. All Recruiting	NGGT INC.	
Gene addition	NCT05972629	IV	Adults (18-65)	SAR444836: AAV carrying cDNA	Phase I/II. Active, not yet recruiting	Sanofi	
Gene editing "phEdit"	NCT05222178	IV	Adult (18-55)	HMN-103: AAVHSC15 containing codon optimised PAH flanked by homology arms targeting PAH intron 1	Phase I. **Discontinued** October 2023	Homology Medicine Inc	(191)

To date, only a limited number of preclinical studies has explored AAV-mediated CRISPR/Cas9 gene editing for PKU, with early work primarily focused on HDR. Richards et al., 2020 (119) employed a dual AAV8 strategy in *Pah*$^{\text{enu2}}$ mice, delivering SpCas9 alongside a donor cassette containing human *PAH* flanked by homology arms to promote targeted integration. Neonatal mice were treated at 1 week of age, with a second dose administered at 5 weeks. Researchers observed only modest reductions in blood phenylalanine, even with the use of a pharmacological inhibitor of NHEJ vanillin. Mean liver PAH activity was 9.6% (SD 1.9%; range 3.3%-24.8%), and blood phenylalanine remained above the therapeutic threshold of 360 µmol/L (~685 µmol/L). These findings highlighted the intrinsic inefficiency of HDR in post-mitotic hepatocytes.

Subsequent work has shifted towards on base and prime editing. The first demonstration came from Villiger et al., (2018) (192), who used a dual AAV8 cytidine base editor in treat *Pah*$^{\text{enu2}}$ mice. They achieved an average of 18.6% allele correction and ~34% wildtype Pah mRNA, reducing blood phenylalanine to physiological levels (120 µmol/L) alongside modest wildtype enzymatic activity (~7-22%) at 8 weeks. Importantly, they showed that as little as 9% allele correction was sufficient to lower blood phenylalanine to below the therapeutic threshold (360 µmol/L).

This work was expanded in 2021, using an optimised base editor delivered via both dual AAV8 and AAV-LNP hybrid system, achieving similar correction efficiencies(~28% dual AAV and ~21% AAV-LNP delivery) and showing no detectable genome- and transcriptome-wide off-target effects (193). More recently, Brooks and colleagues (2024) (194) reported efficient correction of a humanised P281L PKU mouse model using an adenosine base editor delivered as mRNA in LNPs alone. A single administration yielded 28-46% mutation

`

correction without detectable liver toxicity, accompanied by a 36% reduction in blood phenylalanine at 24 hours and 90% reduction after 48 hours treatment, reaching phenylalanine levels to the physiological range.

Prime editing has also been investigated, spearheaded by the Schwank's group in Zürich. In 2022, Bock et al. (195) delivered prime editing reagents to *Pah*$^{\mathrm{enu2}}$ mice using an adenoviral vector and demonstrated an average of 11% correction in neonatal animals, which was sufficient to normalise blood phenylalanine levels to <120 µmol/L. Adult mice showed only a 2% correction and did not meaningful reduce blood phenylalanine (~1000 µmol/L). Building on this work, Rothgangl et al. (196) reported an improved prime editing system in adult PKU mice. Using a hybrid delivery strategy in which AAV8 carried the guide RNAs and LNPs delivered prime editor mRNA, they achieved up to 20% allele correction and reductions in blood phenylalanine to below the therapeutic threshold of 360 µmol/L. A fully LNP-based regimen produced ~8% correction, which also lowered blood phenylalanine into the sub-therapeutic range (<600 µmol/L), though not to normal levels (<120 µmol/L). Collectively, these findings highlight the promise of prime editing but also its current limitations in efficiency and consistency, particularly in adult animals.

In May 2025, Tessera Therapeutics announced the development of a prime editor-like system, "Gene Writer", delivered in LNPs to treat PKU. In non-human primates, Gene Writer achieved ~70% on-target editing at the *PAH* locus in hepatocytes with 67% edited mRNA transcripts, while in a humanised R408W PKU mouse model, ~65% correction of hepatocytes rescued hyperphenylalaninemia (197). While these data are highly promising, they remain unpublished, and the precise mechanism of Gene Writers' action has not yet been disclosed.

`

Table 1.3: Preclinical gene editing studies to treat PKU.

Group/Year	Editing Approach	Vector	Serum Phenylalanine Healthy: <360 μmol/L Physiological: <120 μmol/L	%Threshold to rescue	% WT PAH enzyme activity	Refs
Chen 2020	Nuclease-free HDR	AAV	<360 μmol/L	6% allele correction	No data	(191)
Richards 2020	CRISPR/SpCas9 HDR	AAV	685 ± 82 μmol/L (range: 252-1168 μmol/L)	~10% allele correction; 7.54% ± 2.33% WT PAH mRNA (range: 1.94%–24.00%)	9.6% ± 1.9% (range: 3.3 - 24.8%)	(119)
Villiger 2018	Cytidine base editor	Dual AAV (split intein)	<120 μmol/L Data shown for 8 weeks.	18.6% allele correction; 34.4% WT PAH mRNA	7-22%	(192)
Villiger 2021	Cytidine base editor	Dual AAV (split intein)	<120 μmol/L	28% allele correction	No data	(193)
Villiger 2021	Cytidine base editor	LNP	<360 μmol/L	21% allele correction	No data	(193)
Brooks 2024 †	Adenosine base editor	LNP	< 120 μmol/L	28-46% allele correction	No data	(194)
Böck 2022	Prime editing	AdV	Adults: 1000 μmol/L Neonates: <120 μmol/L	Adults: 2.0 ± 0.2% allele correction Neonatal: 11.1 ± 3.3% allele correction	Adults: No data Neonates: 2-6%	(195)
Rothgangl 2025	Prime editing	Dual AAV-LNP (LNP-PE mRNA, AAV pegRNA)	Variable, ~600 μmol/L (Range 360-900 μmol/L)	20.7% allele correction	2.5-14%, although not explicitly stated	(196)
Rothgangl 2025	Prime editing	LNP	<600 μmol/L	8% allele correction	No data	(196)

All studies using AAV vectors used the AAV8 capsid that were administered using I.V injection

† All studies used Pah^{enu2} mice unless otherwise stated. Brooks et al., (2023) used P281L mouse.

`

1.9 Limitations and future directions

Taken together, gene editing is the best long term treatment option for PKU, able to address both the paediatric and adult populations. However genetic heterogeneity of the disease compilates thing, excluding base and prime editing approaches which can modify single nucleotides or introduce short oligos (100 bp) given practical challenges getting to the clinic. CRISPR/Cas9-mediated double strand break approaches provide a mechanism to develop a universal treatment approach. HDR, while enabling faithful DNA repair, is only active during cell division, limiting the number of editable cells at the time of treatment, and has been shown previously to underperform both preclinical and clinical studies. Homology directed targeted integration (HITI) approaches, which take advantage of the NHEJ constitutively active DNA repair pathway, represent an opportunity to introduce large therapeutic cassettes for the treatment of many pathogenic variants for PKU.

1.10 Requirements for success in the case of PKU

While PAH is expressed in the liver, PKU is not associated with intrinsic liver pathology, making it the ideal target for AAV gene therapy approach. Compared to diseased tissue, a healthy liver allows more consistent vector delivery and transgene expression, as fibrosis, inflammation and altered blood flow can hinder AAV update (165). Additional advantages include high blood flow, fenestrated sinusoidal endothelium that permits vector entry and natural hepatotropic of certain AAV serotypes (17). The common goal for therapeutic interventions for PKU is the restoration of sufficient PAH activity to reduce serum phenylalanine to a clinically safe range (<360 µmol/L) or ideally to physiological levels (<120 µmol/L) without reliance on dietary restriction.

`

PKU is considered a haplosufficient disorder; however, liver biopsy studies of individuals heterozygous for PAH mutations demonstrate residual PAH enzymatic activity ranging from approximately 7-42% of normal across three independent cohorts (198–200). Although these studies employed older methodologies and absolute values should be interpreted with caution, they nonetheless suggest that the therapeutic threshold may be lower than initially assumed. They also highlight the potential for mutated alleles to exert differential effects on wildtype protein, an important consideration again considering the heterogeneity of the condition.

Cell transplantation studies show repopulation of *Pah*$^{\mathrm{enu2/enu2}}$ mouse livers with >10% wildtype cells complete corrects blood phenylalanine levels (201). Notably, similar correction was achieved when engrafting heterozygous hepatocytes, indicating that the limiting factor is the proportion of PAH-expressing cells rather than the total enzymatic activity per cell (202). This is consistent with PAH being a cytosolic enzyme whose expression may be influenced by liver zonation.

Preclinical gene therapy studies using a variety of different approaches (Table 3) demonstrated variable levels of allele correction, wildtype PAH mRNA and wildtype PAH enzyme activity required to rescue blood phenylalanine. Interpretation is complicated, however, by a publication bias towards successful strategies, with few reports describing insufficient or failed approaches to elucidate a true minimum threshold. Taken together, these data suggest that the restoration of >10% wildtype PAH enzymatic activity is generally required to reduce serum phenylalanine to physiological levels.

Looking at preclinical gene editing studies for PKU, the absolute number varies wildly across approaches with as high as 22% activity suggested in recent base editing approaches using both AdV and AAV vectors (147). In the widely used $Pah^{enu2/\mathrm{enu2}}$ mouse model, this corresponds to expression approximately 30% wildtype mRNA levels (147,167). However, the specific mutant alleles involved are likely to influence the requirements for therapeutic success, as different mutations affect PAH tetramer formation, stability, and enzymatic activity to varying degrees. The pathogenic mutation in the $Pah^{enu2/\mathrm{enu2}}$ mouse model is known to exerts a dominant negative effect, forming mixed tetramers with wildtype subunits and reducing overall enzymatic activity (73). Heterozygous unaffected $Pah^{+/\mathrm{enu2}}$ mice display significantly lower PAH activity (25-29%) than expected (50%) (75). There are more than 20 other known pathogenic *PAH* mutations identified (168); however, the larger prevalence of similar dominant-negative variants is currently unknown but an important consideration when considering therapeutic thresholds.

Interpretation is further complicated by a publication bias towards successful strategies, with few reports describing insufficient or intermediary approaches to elucidate a true minimum threshold. Taken together, these data suggest that the restoration of >10% wildtype PAH enzymatic activity is generally required to reduce serum phenylalanine to physiological levels.

1.11 Hypotheses and Aims

Previous attempts using HDR strategies, such as clinical studies by Homology Medicines without the use of CRISPR/Cas9 endonuclease and preclinical studies using CRISPR/Cas9 have been unsuccessful, likely due to insufficient targeted integration of therapeutic cassette. No study has previously applied a HITI approach, DNA repair that is constitutively active, to introduce full-length PAH cDNA for therapeutic benefit.

We hypothesised that dual-AAV delivered CRISPR/Cas9 gene editing tools could be harnessed to achieve mutation-agnostic correction of phenylketonuria (PKU). Specifically, we proposed that targeting *Pah* intron 1 would allow for targeted insertion of a PAH cDNA donor cassette, enable expression of functional PAH protein irrespective of the underlying coding mutation. By doing so, this strategy could provide a universal therapeutic approach using a single set of reagents. This is advantageous when considering future clinical application because safety profiling will only have to be performed once thereby lessening testing as opposed to using unique editing reagents on a mutation-to-mutation basis and ease of manufacturing.

This study aims to:

Aim 1: Develop a dual-AAV CRISPR/Cas9 HITI approach for targeted integration of GFP or PAH donor constructs into *Pah* intron 1.

Aim 2: Evaluate integration efficiency *in vivo* using GFP donors as a proof-of-concept.

Aim 3: Evaluate therapeutic efficacy of the approach using PAH therapeutic donors in a murine model of PKU.

`

2. Materials and Methods

2.1 Materials

All chemicals were handled safely with appropriate personal protective equipment (PPE) and environmental controls in accordance with guidelines outlined in the material safety data sheet (MSDS) of each chemical. Waste products were disposed of safety in concordance with the respective MSDS for each chemical, and internal health and safety policies. Chemicals and biological agents imported from overseas were imported under guidelines from the current (at time of purchase) Permit to import conditionally non-prohibited goods issued to Children's Medical Research Institute by the Australian Government Department of Agriculture, Water and the Environment.

Reagents used to make solution were all analytical grade or higher. All solutions were prepared using MilliQ or molecular biology grade water. Protocols were carried out a room temperature, unless otherwise specified.

2.1.1 Kits and reagents

Table 2.1: Kits and reagents.

Kit or reagent	Cat Number	Source
General		
PureLink DNaseI	12185010	Invitrogen (Waltham, USA)
TurboDNase	AM2238	Invitrogen (Waltham, USA)
Foetal Bovine Serum (FBS)	S-FBS-AU-015	Serana (Brandenburg, Germany)
c0mplete EDTA-free protease Inhibitor cocktail (EDTA-free)	04693132001	Roche (Basel, Switzerland)
Proteinase K	P8107S	New England Biolabs (Ipswich, USA)
PureLink RNase A	12091021	Invitrogen (Waltham, USA)
UltraPure H_20	10977035	Gibco (Waltham, USA)
Tissue Culture		
CyroStor CS10 Cell Freezing Media	12352207	Stem Cell Technologies (Vancouver, Canada)
Dulbecco's Modified Eagle Medium (DMEM)	11965-092	Gibco (Waltham, USA)

`

Dulbecco’s Phosphate Buffer Saline (PBS)	14190144	Gibco (Waltham, USA)
Iscove’s Modified Eagle Medium (IMDM)	I3390-500ML	Sigma-Aldrich (St. Louis, USA)
TrypLE Express Enzyme	12604021	Gibco (Waltham, USA)
Trypan-Blue Solution 4% (w/v)	K940-100ML	Amresco (Dallas, USA)
Penicillin-Streptomycin (PenStrep)	15140122	Gibco (Waltham, USA)
Percoll	17-0891-01	GE Healthcare (Chicago, USA)
Molecular Biology		
Restriction enzymes + corresponding buffers	Various	New England Biolabs (Ipswich, USA)
DNA Loading Buffer 6×	B7024S	New England Biolabs (Ipswich, USA)
dNTPs 40mM	BIO-39043	Bioline (London, UK)
Isolate II Plasmid Mini kit	BIO-52057	Bioline (London, UK)
HyperLadder 1kb	BIO-33026	Bioline (London, UK)
HyperLadder 50bp	BIO-33040	Bioline (London, UK)
PGEM-T Easy Vector System	A1360	Promega (Maddison, USA)
NucleoBond Xtra Maxi Plasmid DNA Purification kit	740416.50	Macherey-Nagel (Düren, Germany)
PureLink RNA Mini kit	12183025	Invitrogen (Waltham, USA)
Qubit BR dsDNA Assay kit	Q33265	Invitrogen (Waltham, USA)
Redsafe Nucleic Acid Staining Solution	21141	iNtRON Biotechnology (Gyeonggi-do, South Korea)
T4 DNA Ligase + Buffer	M0202S	New England Biolabs (Ipswich, USA)
T4 DNA polymerase + Buffer	M0203S	New England Biolabs (Ipswich, USA)
T4 Polynucleotide Kinase (PNK) enzyme + Buffer	B0201S	New England Biolabs (Ipswich, USA)
Wizard SV Gel and PCR Clean-Up System	A9282	Promega (Maddison, USA)
Monarch Spin gDNA Extraction Kit	T3010	New England Biolabs (Ipswich, USA)
PCR		
Q5 High-Fidelity DNA Polymerase + Buffer	M0491	New England Biolabs (Ipswich, USA)
SuperScript IV First-Strand Synthesis Kit	18091050	Invitrogen (Waltham, USA)
ddPCR EvaGreen supermix	186-4035	Bio-Rad (Hercules, USA)
KAPA PROBE FAST qPCR Mastermix	07959818001	KAPA Biosystems (Wilmington, USA)
2 × TB Green Premix Ex Taq	RR82WR	TAKARA (Kyoto, Japan)
Viral production		
Benzonase Endonuclease	E1014	Sigma-Aldrich (St. Louis, USA)
Animal Studies		
NeoBase 2 Non-derivatised MSMS kit	3044-0010	Revvity (Waltham, USA)
Isoflurane		Abbott Australia (Macquarie Park, Australia)
Standard Chow - 20% (w/w) protein	SF00-100	Speciality Feeds (Glen Forrest, Australia)
High Protein Chow - 61% (w/w) protein	SF18-142	Speciality Feeds (Glen Forrest, Australia)
PAH *in vitro* Enzymatic Assay		
DC Protein Assay Kit	500-0007	Bio-Rad (Hercules, USA)
Catalase	30-100MG	Sigma-Aldrich (St. Louis, USA)

Tetrahydrobiopterin (BH4)	11.212-1	Schirck Laboratories (Bauma, Switzerland)
Liver perfusion		
HBSS -/-	H6648	Sigma-Aldrich (St. Louis, USA)
Collagenase IV powder	C5138-1G	Sigma-Aldrich (St. Louis, USA)
DNaseI	DN25-100mg	Sigma-Aldrich (St. Louis, USA)
Liver perfusion Kit (mouse and rat)	130-128-030	Miltenyi Biotec (Bergisch Gladbach, Germany)
MACS tissue storage solution	130-100-008	Miltenyi Biotec (Bergisch Gladbach, Germany)
Immunohistochemistry		
4’-6-Diamidino-2-phenylindole dihydrochloride (DAPI)	D9542	Sigma-Aldrich (St. Louis, USA)
Donkey serum	D9663	Sigma-Aldrich (St. Louis, USA)
Goat serum	G9023	Sigma-Aldrich (St. Louis, USA)
Tissue-Tek Optimal Cutting Temperature (O.C.T) Compound	4583	Sakura (Osaka, Japan)

2.1.2 Buffers and Solutions

Table 2.2: Buffers and Solutions.

Solution	**Composition**	**Preparation and storage**
General		
4% (w/v) paraformaldehyde (PFA)	4% (w/v) paraformaldehyde (PFA) in 1 × PBS	8 g Paraformaldehyde in 100 mL, dissolved at 60ºC, in fume cabinet. Raise pH with NaOH until dissolved (pH 7.2-7.4). Made up to 200 mL with 2 × PBS. Stored at -20ºC.
1 × Phosphate Buffered Saline (PBS)	1 × PBS (no Ca^{+2}, no Mg^{+2})	1 × PBS tablet per 500 mL milliQ water
10% TE	1 mM Tris, 0.1 mM EDTA	0.5 mL 1 M Tris, 0.1 mL 0.5 M EDTA. MilliQ to 500 mL. Adjust to pH 8 and autoclave. Keep sterile.
Tissue culture		
Complete DMEM or Growth media	DMEM + 10% (v/v) heat-inactivated FBS + 1 × penicillin streptomycin antibiotic solution	50 mL FBS, 5 mL PenStrep to 445 mL DMEM. Filter through 0.22 µM membrane and keep sterile. Store at 4°C.
Maintenance media	DMEM + 2% (v/v) heat-inactivated FBS + 1 × penicillin streptomycin antibiotic solution	10 mL FBS, 5 mL PenStrep to 485 mL DMEM. Filter through 0.22 µM membrane and keep sterile. Store at 4°C.
Transfection media	IMDM + 10% (v/v) heat-inactivated FBS + 1 × glycine + 1 × penicillin streptomycin antibiotic solution	50 mL FBS, 5 mL glycine, 5 mL PenStrep to 440 mL IMDM. Filter through 0.22 µM membrane and keep sterile. Store at 4°C.
Molecular Biology		
Agarose gel	0.5%-2% (w/v) agarose dissolved in 1 × TBE.	Agar weighed and dissolved in 1× TBE buffer. Microwaved on low until completely dissolved. Set at room temperature.

Luria-Bertani (LB) agar	LB broth containing 1.5% (w/v) agar.	4.5 g agar to 300 mL LB broth. Autoclaved on the same day then store at RT.
Luria-Bertani (LB) broth	1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl.	10 g Tryptone, 5 g yeast extract, and 10 g NaCl per 1 L milliQ water. Autoclaved on the same day then stored at RT.
Liver Lysis Buffer	10 mM Tris-HCl, 0.1 M EDTA, 0.5% (v/v) SDS	100 mL 500 mM EDTA, 5 mL 1M EDTA, 25 mL 10% SDS to up to 500 mL water.
5 × TBE	0.445 M Tris, 0.445 M Borate, 0.01 M EDTA	108 g Tris base, 55 g Boric acid, 40 mL 0.5 M EDTA pH 8.0 in 2L milliQ water
1 × TBE	89 mM Tris, 89 mM Borate, 2mM EDTA	Dilute 5 × TBE 1 in 5 with milliQ water.
1000 × MgCl2/CaCl2 solution	0.901 M $CaCl_2$, 0.493 M $MgCl_2$	66.15 g CaCl2.2H20 + MgCl2.6H20 up to 500 mL milliQ. Filter sterilise (0.22 µM).
PCR		
ddPCR Alkaline digestion buffer	25 mM NaOH, 0.2 mM EDTA	1.25 mL of 1 M NaOH, 20 µL of 0.5 M EDTA, 48.73 mL H_2O. Store at RT.
ddPCR Neutralisation buffer	40 mM Tris, 0.05% (v/v) Tween20	2 mL of 1 M Tris, 25 µL of Tween 20, make up to 50 mL with H_2O, pH 5, stored at 4°C.
Viral Manufacture		
Viral Production		
Benzonase Buffer	50 mM Tris, 2 mM $MgCl_2$	20 mL 1 M Tris-HCl, pH 8.5, 1 mL 1 M $MgCl_2$, milliQ water to 500 mL. Autoclaved.
2 M $CaCl_2$	2 M $CaCl_2$	29.404 g $CaCl_2$.2H2O in 100 mL milliQ Water. Autoclaved
1 M $CaCl_2$	1 M $CaCl_2$	14.702 g $CaCl_2$.2H2O in 100 mL milliQ Water. Autoclaved
0.15 M Na_2HPO_4	0.15 M Na_2HPO_4	10.647 g Na_2HPO_4 in up to 500 mL dH2O, pH 7.10. Autoclaved.
2 × HEPES-buffered Saline (HBS)	280 mM NaCl, 50 mM HEPES	8.182 g NaCl, 5.958 g HEPES in up to 500 mL milliQ. Adjust to pH 7.10. Autoclaved.
HEPES/EDTA resuspension buffer	50 mM HEPES, 0.15 M NaCl, 25 mM EDTA	Adjust to pH 7.4. Filter sterilise through 0.22 mM membrane and store at RT.
Transfection solution A (per 15cm tissue culture dish)	32 µg Plasmid DNA, 287.5 µL 2 M $CaCl_2$ made up to 2.5mL 10% (v/v) TE	
Transfection solution B (per 15cm tissue culture dish)	23 mL 0.15 M Na_2HPO_4, 2277 µL 2× HBS	
Viral purification via Caesium chloride gradients		
Caesium chloride solution (1.3 g/mL)		39.9 g to 100 mL PBS (Ca^{2+}/Mg^{2+}). Check density and filter sterilise (0.22 µM).
Caesium chloride solution (1.37 g/mL)		50 g to 100 mL PBS (Ca^{2+}/Mg^{2+}). Check density and filter sterilise (0.22 µM).
Caesium chloride solution (1.5 g/mL)		33.75 g to 50 mL PBS (Ca^{2+}/Mg^{2+}). Check density and filter sterilise (0.22 µM).
Dialysis Buffer	1× PBS (Ca^{2+}/Mg^{2+}) ± 5% (v/v) glycerol	4 L 1 × PBS + 4mL 1000 × MgCl2/CaCl2 solution + 200 mL glycerol
PAH *in vitro* Enzymatic Assay		

Thöny lab protein extraction buffer (203)	50mM Tris-HCl, 0.1M KCl, 1mM EDTA, 1mM DDT, c0mplete protease inhibitor cocktail tablet (EDTA-free version)	pH 7.5. Solution stable for 1 week at 4C as c0mplete protease inhibitor cocktail unstable.
1M KCl		7.5 g KCl in 100 mL water. Filter sterilise.
2× Substrate solution	150 mM potassium chloride, 400 µg/mL catalase, 300 mM ammonium iron(II) sulfate, 2 mM dithiothreitol, 200 mM phenylalanine-d5, and 200 µM tetrahydrobiopterin (BH4)	Solution made immediately before use as BH4 is very unstable.
2× Enzyme solution	150 mM potassium chloride, 400 µg/mL catalase, 300 mM ammonium iron (II) sulfate, 2 mM dithiothreitol, and 225 µg/mL liver extract	
Immunohistochemistry		
10% (w/v) sucrose solution		20 g in 200 mL 1 × PBS. Filter sterilised (0.22µM).
20% (w/v) sucrose solution		40 g in 200 mL 1 × PBS. Filter sterilised (0.22 µM).
30% (w/v) sucrose solution		60 g in 200 mL 1 × PBS. Filter sterilised (0.22 µM).
Gelvatol mounting medium	13.3% (w/v) PVA, 0.133M Tris HCl	15 mL milliQ water with 30 mL 0.2 M Tris HCl (pH 8.5) and 6 g of polyvinyl alcohol
PBS + 0.1% (v/v) Triton X-100	PBS + 0.1% (v/v) Triton X-100	500 µL Triton X-100 + 500 mL 1 × PBS. Mix well.
PBST	PBS + 0.1% (v/v) Tween20	1 mL Tween20 in 1L 1 × PBS. Mix well.
Donkey blocking buffer	10% donkey serum, 10% FBS in PBS	1 mL donkey serum, 1 mL FBS and 8 mL PBS.
Rabbit blocking buffer	10% rabbit serum, 10% FBS in PBS	1 mL rabbit serum, 1 mL FBS and 8 mL PBS.
Animal studies		
FLOW buffer	5% (v/v) FBS in PBS	
Genotyping		
'HotSHOT' Alkaline digest buffer (204)	5 mM NaOH, 0.2mM disodium EDTA in water	pH 12
'HotSHOT' Neutralisation buffer (204)	40 mM Tris-HCl in water	pH 5

2.1.3 Chemicals

All chemicals required to make solutions in Table 2 were analytical grade or higher.

Table 2.3:Chemicals.

Chemical name	**Source**
L-Phenylalanine (ring-D_5)	Cambridge Isotope Laboratories Inc (Massachusetts, USA)
L-Tyrosine (ring-D_4)	Cambridge Isotope Laboratories Inc (Massachusetts, USA)
Agar	Meridian Bioscience (Cincinnati, USA)
Agarose	Meridian Bioscience (Cincinnati, USA)
Ammonium acetate	Sigma-Aldrich (St. Louis, USA)

Ampicillin Sodium Salt	AG-Scientifics (San Diego, USA)
Ammonium iron (II) sulfate hexahydrate	Sigma-Aldrich (St. Louis, USA)
Boric Acid	Chem Supply (Gillman, Australia)
Calcium Chloride	Sigma-Aldrich (St. Louis, USA)
Caesium Chloride	Invitrogen (Waltham, USA)
Chloroform	VWR (Pennsylvania, USA)
Chloroform:isoamyl alcohol 24:1 (v/v)	Sigma-Aldrich (St. Louis, USA)
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich (St. Louis, USA)
Dithiothreitol (DTT)	Sigma-Aldrich (St. Louis, USA)
Ethylenediaminetetraacetic acid (EDTA)	Chem Supply (Gillman, Australia)
Ethanol 70% (v/v)	POCD Healthcare (New South Wales, Australia)
Ethanol 100% (v/v)	POCD Healthcare (New South Wales, Australia)
Formic acid	Sigma-Aldrich (St, Louis, USA)
Glycerol	Merch (Rahway, USA)
4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)	Astral Scientifics (Sydney, Australia)
Hydrochloric acid	Merck (Rahway, USA)
Isopropanol	Univar (Illinois, USA)
Kanamycin sulfate	Gibco (Waltham, USA)
Magnesium Chloride	Sigma-Aldrich (St. Louis, USA)
Methanol	POCD Healthcare (New South Wales, Australia)
2-methylbutane	Sigma-Aldrich (St. Louis, USA)
Neutral buffered formalin 10% (v/v)	Sigma-Aldrich (St. Louis, USA)
Paraformaldehyde	Merck (Rahway, USA)
Phosphate buffered saline (PBS) tablets	MP Chemicals (Melbourne, Australia)
Polyethylene glycol (PEG) 800	Sigma-Aldrich (St. Louis, USA)
Polyvinyl alcohol	Sigma-Aldrich (St. Louis, USA)
Phenol:chloroform:isoamyl alcohol 25:24:1 (v/v/v)	Sigma-Aldrich (St. Louis, USA)
Sodium Chloride	Sigma-Aldrich (St. Louis, USA)
Sodium dodecyl sulfate (SDS)	Sigma-Aldrich (St. Louis, USA)
Sodium Hydroxide pellets	Sigma-Aldrich (St. Louis, USA)
Sucrose	Chem Supply (Gillman, Australia)
Triton X-100	Sigma-Aldrich (St. Louis, USA)
Trizma Base	Amresco (Dallas, USA)
Tris(hydroxymethyl)aminomethane (Tris)	Chem Supply (Gillman, Australia)
TRIzol	Invitrogen (Waltham, USA)
Tryptone	Sigma-Aldrich (St. Louis, USA)
Tween-20	Sigma-Aldrich (St. Louis, USA)
Yeast Extract	Scharlau (Barcelona, Spain)

2.1.4 Immortalised cell lines

All cell lines used were adherent cells.

Table 2.4: Immortalised mammalian cell lines.

Cell lines	Descriptions	References
HEK293	Human embryonic kidney cells	(205)
NIH3T3	Mouse embryonic fibroblasts	(206)
B35	Rat neuroblastoma cells	(207)

2.1.5 Competent bacterial cells for cloning

Table 2.5: Competent bacterial Escherichia coli strains used for cloning.

Strain	Genotype	Source
DH5-α	F – φ80lacZΔM15 Δ(lacZYA-argF)U169 recA1 endA1 hsdR17(rK –, mK +) phoA supE44 λ– thi-1 gyrA96 relA1	Thermo Scientific (Massachusetts, USA)
SURE	F′ lacIq ZΔM15 Tn10(TetR) Δ(mcrC-mrr)114::IS10 endA1 recB21 recJ21 h sdR17(r_K^– m_K^+) mcrA mcrBC	Agilent (Santa Clara, USA)

2.1.6 Antibodies

Table 2.6: Antibodies.

Antibody	Reactivity	Host	Source	Catalogue #	Clonality	Conjugate	Dilution
Anti-GFP	NA	Goat	Rockland	600141215	Poly-	DyLight488	1:800
Anti- GLUL	NA	Mouse	Origene	CF500700	Mono-	NA	1:800
Alexa-Fluor 594	Rabbit	Donkey	Invitrogen	A21207	Poly-	NA	1:800

NA: Not applicable.

2.1.7 Plasmids

Select plasmid sequence and maps can be found in Appendices Section 6.1.

Table 2.7: Plasmids.

Plasmid Name	Details	Source	Ref
pX601-AAV-CMV::NLS-SaCas9-NLS-3×HA-BGHpA;U6::BsaI-sgRNA	CRISPR/SaCas9 plasmid. SaCas9 under ubiquitous promoter (CMV). SgRNA expressed from ubiquitous promoter (U6). Used for cloning candidate guides and testing performance *in vitro*.	Addgene (Watertown, USA)	(174)
pX602-AAV-TBG::NLS-SaCas9-NLS-3×HA-BGHpA;U6::BsaI-sgRNA	CRISPR/Cas9 plasmid. Cas9 under liver-specific promoter (CMV). SgRNA expressed from ubiquitous promoter (U6). Used for cloning candidate guides and testing performance *in vivo*.	Addgene (Watertown, USA)	(174)

pGEM-T Easy	TA Cloning vector	Promega (Maddison, USA)	
pAAV2/8	Packaging plasmid for AAV8	Dr James Wilson, University of Pennsylvania	(208)
pAd5	Helper virus. Adenovirus type 5 genome.	Addgene (Watertown, USA)	(209)
pX601-AAV-CMV::NLS-SaCas9-NLS-3×HA-BGHpA;U6::BsaI-sgRNA1	AAV CRISPR/SaCas9 plasmid used for *in vitro* guide screening	Created by Caitlin Lucas	(174)
pX601-AAV-TBG::NLS-SaCas9-NLS-3×HA-BGHpA;U6::BsaI-sgRNA2	AAV CRISPR/SaCas9 plasmid used for *in vitro* guide screening	Created by Caitlin Lucas	(174)
pX601-AAV-TBG::NLS-SaCas9-NLS-3×HA-BGHpA;U6::BsaI-sgRNA3	CRISPR/SaCas9 plasmid used for *in vitro* guide screening	Created by Caitlin Lucas	(174)
pX601-AAV-TBG::NLS-SaCas9-NLS-3×HA-BGHpA;U6::BsaI-sgRNA4	AAV CRISPR/SaCas9 plasmid used for *in vitro* guide screening	Created by Caitlin Lucas	(174)
pX602-AAV-TBG::NLS-SaCas9-NLS-3×HA-BGHpA;U6::BsaI-sgRNA1	AAV CRISPR/SaCas9 plasmid used for *in vivo* guide screening	Created by Caitlin Lucas	(174)
pX602-AAV-TBG::NLS-SaCas9-NLS-3×HA-BGHpA;U6::BsaI-sgRNA2	AAV CRISPR/SaCas9 plasmid used for *in vivo* guide screening	Created by Caitlin Lucas	(174)
pX602-AAV-TBG::NLS-SaCas9-NLS-3×HA-BGHpA;U6::BsaI-sgRNA3	AAV CRISPR/SaCas9 plasmid used for *in vivo* guide screening	Created by Caitlin Lucas	(174)
pX602-AAV-TBG::NLS-SaCas9-NLS-3×HA-BGHpA;U6::BsaI-sgRNA4	AAV CRISPR/SaCas9 plasmid used for *in vivo* guide screening	Created by Caitlin Lucas	(174)
pAAV2-LPhOTC BB2	scAAV backbone used to generate donor vectors. Cloning.	A/Prof Samantha Ginn	(210)
pUC57-Kan-sgRNA1-splice GFP-sgRNA1	Cloning plasmid with eGFP reporter downstream of splice acceptor site. Used to generate donor vectors.	GenScript (Nanjing, China)	
pUC57-Kan-sgRNA1-2PA GFP-sgRNA1	Cloning plasmid with eGFP reporter downstream of 2PA self-cleaving peptide sequence. Used to generate donor vectors.	GenScript (Nanjing, China)	
pUC57-Kan-sgRNA1-Splice hcoPAH-sgRNA1	Cloning plasmid containing partial the coding sequence of human PAH (exons 2-13) downstream of splice acceptor site. Used to generate donor vectors.	GenScript (Nanjing, China)	
pUC57-Kan-sgRNA1-2PA hcoPAH-sgRNA1	Cloning plasmid with full coding sequence of human PAH (exons 1-13) reporter downstream of 2PA self-cleaving peptide sequence. Used to generate donor vectors.	GenScript (Nanjing, China)	
pAAV2-Splice GFP BB2	AAV HITI donor plasmid	Created by Caitlin Lucas	
pAAV2-2PA GFP BB2	AAV HITI donor plasmid	Created by Caitlin Lucas	
pAAV2-Splice PAH BB2	AAV HITI donor plasmid	Created by Caitlin Lucas	
pAAV2-2PA PAH BB2	AAV HITI donor plasmid	Created by Caitlin Lucas	

2.1.8 PCR primers

Table 2.8: PCR primers.

Description	**Direction**	**Name**	**Sequence 5' - 3'**
General PCR			
Genotyping	Sense	CL13F	CTGTTGCTGGCTTACTGTCG
	Antisense	CL14R	TGTGTACATGGGCTTAGATCC
Indel PCR			
Indel PCR for sgRNA1	Sense	CL87F	AGCTCTGGGTTTCTATTACAGGA
	Antisense	CL58R	TGAGCCATGACCTACTTTGGT
Indel PCR for sgRNA3	Sense	CL65F	GCTTTTGTGAGGACTGAATGGT
	Antisense	CL66R	AGAGCCAGGTTGTTTTCTGT
Indel PCR for sgRNA2 and sgRNA4	Sense	CL59F	GTGTGTCAATCCACATGCCTG
	Antisense	CL60R	GACACTGTCCTGATCTGCCC
Junction analysis dPCR			
Genome-GFP insert junction analysis	Sense	CL87F	AGCTCTGGGTTTCTATTACAGGA
	Variable depending on insert orientation	SG76R	GGTGCAGATGAACTTCAGGGT
	Variable depending on insert orientation	SG408F	GCTTCTGAGGCGGAAAGAAC
	Antisense	CL58R	TGAGCCATGACCTACTTTGGT
Genome-PAH insert junction analysis	Sense	CL87F	AGCTCTGGGTTTCTATTACAGGA
	Variable depending on insert orientation	CL68R	GACAGGATTTGGTTGGCA
	Variable depending on insert orientation	SG408F	GCTTCTGAGGCGGAAAGAAC
	Antisense	CL58R	TGAGCCATGACCTACTTTGGT
Vector titre			
BGH poly A (BGHpA)	Sense	BGHpA_F	GCCTTCCTTGACCCTGGA
	Antisense	BGHpA_R	ACTCAGACAATGCGATGCAA
Vector copy number (VCN)			
mAlbumin	Sense	mALB_F	AACTGCTACTCCCCTCCTAC
	Antisense	mALB_R	TTTACCCCAGTGCAGGAAAG
GFP	Sense	GFP_F	TCAAGATCCGCCACAACATC
	Antisense	GFP_R	TTCTCGTTGGGGTCTTTGCT
cohPAH vector	Sense	CL300F	TGCAGCTACGATCCCTAGAC
	Antisense	CL301R	AGATCAAGTGAGATATCGTCGAC
SaCas9	Sense	330_F	GATTCGACGTGTACCTGGAC
	Antisense	331_R	CGTTGTTGTAGAAGGAGGCG
dPCR transcript			
Hprt transcript (150)	Sense	Hprt_F	GTTAAGCAGTACAGCCCCAAAATG
	Antisense	Hprt_R	AAATCCAACAAAGTCTGGCCTGTA
Murine *Pah* transcript	Sense	CL211F	AGCTGTTGTCCTGGAGAACG
	Antisense	CL115R	GGAGTTGTCTTCGATGTAACTT
Hybrid GFP transcript	Sense	CL211F	AGCTGTTGTCCTGGAGAACG

	Antisense	SG633R	TCCAGCTCGACCAGGATG
Hybrid hcoPAH transcript	Sense	CL211F	AGCTGTTGTCCTGGAGAACG
	Antisense	CL119R	GATTCAATGTGTGTCAGGTTGATC

2.1.9 PCR probes

Probes purchased from IDT (Coralville, USA).

Table 2.9: PCR probes.

Description	Name	Sequence 5' - 3'	Details
Genotyping probe for wildtype allele	WT-PKU	TCCGAGTCTTCCACTGCAC	/5HEX/TC CGA GTC T/ZEN/T CCA CTG CAC /3IABkFQ/
Genotyping probe for mutant allele	MUT-PKU	TCCGAGTCTCCCACTGCAC	6-FAM™/ZEN™/3' IB®FQ

2.2 Tissue Culture

Cell culture was performed using aseptic technique and sterile equipment within a Class II biohazard safety cabinet. Cells were maintained in disposable plastic flasks and dishes (Corning) and maintained at 37°C in 5% $C0_2$ humidified incubator. Growth media and maintenance media were supplemented with 2-10% (v/v) heat-inactivated foetal bovine serum (FBS). FBS was heat-inactivated at 56°C for 30 minutes prior to use to inactivate complement proteins and minimise batch-to-batch serum variability. FBS was aliquoted and stored at -20°C until use. Sterile 1× PBS (-/-) was used to wash cells and 1× TrypLE Express Enzyme used to dissociate cells. Tissue culture reagents were pre-warmed to 37° C in a water bath prior to use.

2.2.1 Routine culturing and cell counting

Adherent cells were passaged by first aspirating old cell media, washing with 1× PBS (-/-) and dissociated with 4 mL 1×TrypLE Express Enzyme for 10 minutes at 37° C. After incubation, flasks were gently tapped to aid in cell detachment and promote a single cell suspension. TrypLE was neutralised with 7 mL DMEM, and the cell suspensions were combined and mixed well for equal dispersion. Suspended cells were then added to new flasks at the desired density (dependent on cell type and downstream application) and the appropriate volume of 10% DMEM added for flask size (20-25 mL T150 flasks or 20 mL for 15cm dishes).

Cells were counted by diluting 1:5 or 1:10 Trypan Blue, incubating for 2 minutes and applying 10 µL to hemocytometer. Counts were performed either manually using a standard hemocytometer or with an automated hemocytometer (Countess 3FL, Invitrogen). Only viable cells (those that did not take up the dye) were included in the final count. Cell seeding density varied depending on the cell type used and the timing of the next planned passage, with the aim of reaching approximately 80% confluency at the time of splitting.

2.2.2 Cell recovery from cryostorage

Frozen cells were rapidly thawed at 37°C in a water bath, followed by the gradual addition of 11 mL of 10% DMEM to minimise osmotic shock. Cell suspension was centrifuged at 40×*g* for 3 minutes, the supernatant discarded, and the cell pellet resuspended in fresh 10% DMEM. Cells were then seeded in a T75 flask and passaged at least three times prior to experimental use.

2.2.3 Cryopreservation

Dissociated cells were pelleted by centrifugation at 1500×*g* for 5 minutes, then resuspended in growth media to a density of 1-5×10^6 cells/mL. Cell suspension was aliquoted into Cyrofreezer internal thread tube then placed in CoolCell LX Cell Freezing container

(Corning) and stored in -80°C. Cells were stored short term in -80°C (up to 2 months) or transferred to liquid nitrogen for long term storage.

2.3 Molecular biology

2.3.1 DNA extraction from cells

Genomic DNA was extracted from cells using the Monarch spin gDNA extraction kit (New England Biolabs) according to manufacturer's instructions. Briefly, cells (1×10^4 to 5×10^6) were pellet by centrifugation at 1500rpm, media removed and stored at -20°C until extraction.

Frozen cells were thawed on ice, resuspended in 100 µL cold PBS with 1 µL Proteinase K and 3 µL RNase A and lysed in 100 µL cell lysis buffer. After a 5-minute incubation at 56°C, 400 µL of gDNA binding buffer was added and samples were vortexed briefly. Lysates were loaded into spin columns and centrifuged to bind and purify DNA, followed by two wash steps. DNA was eluted in 100 µL of preheated elution buffer and quantified using Nanodrop spectrophotometer (Section 2.3.10).

2.3.2 Restriction enzyme digest

Restriction enzyme digestion was performed using restriction enzymes from New England Biolabs. Reactions were performed as per manufacturer instructions with the recommended enzyme buffer for highest enzyme activity. Restriction enzyme used is indicated where appropriate.

`

2.3.3 Agarose gel electrophoresis

Agarose gel electrophoresis was used to separate DNA fragments according to molecular weight. Gels were prepared by dissolving 0.5-2% (w/v) agarose in 1×TBE through heating. RedSafe Nucleic Acid Staining solution (Bioline) was added at a ratio of 5 μL per 100 mL of gel solution prior to casting. Molten gel was poured into a casting tray with combs inserted and allowed to set at room temperature prior to loading samples. DNA loading buffer (6×) was added to samples prior to loading into gel wells. Hyperladder 1kb and Hyperladder 50 bp (Bioline) ladders were routinely used as molecular weight markers. Electrophoreses was typically initiated at 60V for 5 minutes to allow samples to enter the gel, followed by a voltage increase to 100-120V for 30 minutes to 1.5 hours. Gels were run using Bio-Rad PowerPac Universal Power packs and visualised using AXYGEN Gel Documentation System.

2.3.4 Purification of DNA from agarose gel slices

If recovery of DNA was required, the desired DNA band was visualised under a UV transilluminator and excised from the gel using a sterile scalpel. Gel slices were transferred into 15 mL Falcon tubes and stored at –20°C until further processing.

DNA purification from the excised gel bands was performed using the Wizard SV Gel and PCR Clean-Up System (Promega) as per manufacturer instructions. Briefly, gel slices were dissolved at 50°C for approximately 10 minutes in membrane binding solution (1 μL per 1 mg of gel) until fully dissolved. The solution was then loaded onto spin columns, washed twice and DNA eluted in 30 μL nuclease-free water.

`

2.3.5 A-tailing of PCR DNA fragments

A-tailing refers to the addition of a single adenine base to the 3' termini of a DNA fragment. This was performed on DNA fragments generated using the Q5 high fidelity polymerase (New England Biolabs) which produces blunt ends, to facilitate downstream cloning. Reactions were performed using Taq DNA polymerase and ThermoPol Buffer (New England Biolabs) and consisted of 7 µL column-purified PCR product (purified as per section 2.3.3), 1×ThermoPol Buffer, 1 µM dATP and 5 units of Taq DNA polymerase and incubated in a thermocycler at 37°C for 1 hour.

2.3.6 DNA ligation

DNA ligations reactions were performed to join fragments of DNA one or more ends. Reactions were performed using T4 DNA ligase kit (New England Biolabs) according to manufacturer's instructions. Each ligation reaction consisted of 1 µL of pre-cut plasmid backbone DNA (100ng/ µL), 7 µL of A-tailed insert, 2 µL of 10× T4 DNA ligase buffer and 1 µL of T4 DNA ligase. The insert and vector were combined at a molar ratio of approximately 3:1 (insert:vector) to enhance ligation efficiency using NEBbiocalculator (New England Biolabs, 2024). Reactions were incubated overnight at 4°C or at 15°C for at least 4 hours.

2.3.7 Transformation of competent cells

Competent cells (Table 5) were thawed on ice until just thawed. A 30 µL aliquot of cells was transferred to a pre-chilled 14 mL polypropylene round bottle tube. To each tube, 3 µL of DNA (either ligation mix or plasmid) was added and gently mixed by flicking. Tubes were incubated on ice for 20 minutes. Cells were then heat-shocked in a water bath set at 42°C for 45 seconds and immediately returned to ice for 2 minutes. Following heat shock, 950 mL LB broth was added to each tube, and the cells were incubated for 1 hour at 37°C while shaking

`

(225 rpm) to allow recovery and expression of the selective antibiotic resistance gene. After recovery, cells were plated on LB agar plates containing appropriate selection antibiotic (Ampicillin at 100ug/mL or Kanamycin 50ug/mL) at various volumes (50 µL, 200 µL and the remaining volume). Plates were allowed to dry and incubated inverted overnight at 37°C.

2.3.8 Sequencing of cloned products

Sanger sequencing was outsourced to the Australian Genome Research Facility (AGRF) Sydney, Australia. Reactions contained 1 µL of 10 µM sequencing primer and the appropriate amount of DNA (600-1200 ng for double-stranded plasmid, 3-8 ng for 100-200 bp PCR products, 6-12 ng for 200-400 bp PCR products, 12-18 ng for 400-600 bp PCR product, 18-30 ng for 600-800 bp PCR products), made up to a final volume of 12 µL with nuclease-free water.

2.3.9 Propagation and isolation of plasmid DNA

2.3.9.1 Small-scale plasmid purification

Small scale plasmid DNA isolation from bacterial clones was performed using Isolate II Plasmid Mini kit (Bioline). Single bacteria colonies were selected from either LB/ampicillin or LB/Kanamycin agar plates. Colonies were inoculated in 3 mL LB broth containing selective antibiotic (ampicillin 100 µg/mL or kanamycin 50 µg/mL) in 15 mL loosely capped, round bottom polypropylene tubes. Cultures were incubated overnight (~16 hours) at 37°C on orbital shaker (225 rpm). 1 mL of cultures were purified using Isolate II Plasmid Mini kit (Bioline) as per manufacturer’s instructions while the other 2 mL stored at 4°C until further use. Plasmid DNA subject to restriction enzyme digest (Section 2.3.1) and Sanger sequencing (Section 2.3.8) to ensure product of desired sequence.

2.3.9.2 Large-scale plasmid purification

If larger scale plasmid production was required, after confirming cloned product following a miniprep, the remaining 2mL of culture was introduced into 500 mL conical flasks with 300 mL LB, 3 mL selective antibiotic and were incubated overnight at 37°C on an orbital shaker at 225 rpm. Cultures were centrifuged at 7500 rpm for 15 minutes to pellet cells that were stored at -20°C until purification. Large bacterial cultures (expected yield >2000 µg) were purified using Nucleobond Xtra Maxi kit (Macherey Nagel) as per manufacturer's instructions. Purified maxiprep sequence was confirmed using both Sanger sequencing and restriction enzyme digestion. Restriction enzyme digest was important for verifying the integrity of AAV plasmids, as they contain inverted terminal repeats (ITR) which are highly GC-rich repetitive elements that form secondary structures that are difficult to assess by Sanger sequencing. ITRs are prone to instability during bacterial propagation, leading to partial deletions or truncations that can impair viral packaging efficiency. Restriction digests therefore provided a reliable method to confirm the presence of full-length ITR sequences prior to downstream AAV production.

2.3.10 Nucleic acid quantification

Quantification of DNA and RNA was performed using either a Nanodrop Once/OneC Microvolume UV-Vis spectrophotometer (Thermo Scientific) or Qubit 4 Fluorometer (Invitrogen) depending on the application. The Nanodrop was used to quantify DNA for general molecular protocols. Sample purity was assessed by evaluating absorbance ratios of A260/A280 and A260/A230. Although the NanoDrop provides a rapid and cost-effective method of DNA/RNA quantification, it is sensitive to contaminants (proteins, phenol, guanidine and salts etc) which can lead to overestimation of nucleic acid concentrations. For applications requiring precise DNA quantification, such digital droplet PCR (ddPCR) and digital PCR (dPCR), the Qubit 4 Fluorometer was used. DNA concentration was determined using the Qubit dsDNA BR Assay Kit, which uses fluorescent dyes that selectively bind to dsDNA. Samples and standards were prepared according to manufacturer's instructions.

2.4 Amplification of DNA fragments using polymerase chain reaction (PCR)

2.4.1 General PCR

General polymerase chain reaction (PCR) was performed using Q5 High-Fidelity DNA Polymerase (New England Biolabs) and custom-designed primers synthesised either by Sigma or BioNeer (Table 8). PCR reactions were prepared to a final volume of 50 µL, containing 1 Unit Q5 High-Fidelity DNA Polymerase, 1× Q5 High-Fidelity Reaction Buffer, 10 mM of each dNTP, 25 µM of each forward and reverse primer, template DNA (3 µL of variable concentration) and nuclease-free water to volume. Primers used were specified in Table 8.

PCRs were run on Eppendorf Mastercycler X50 with a constant ramp rate of 5°C/second. PCR conditions were set as an initial denaturation step of 98°C for 30 seconds, followed by 40 cycles of denaturation (98°C for 15 seconds), variable annealing temperature (20 seconds) and extension (72°C for typically 15 seconds). Extension time was typically 15 seconds, however, this was dependent on the expected product size, using an estimated rate of 20-30 seconds per kilobase of DNA. A final step was performed at 72°C, after which the reactions were held at 10°C. Annealing temperature was determined for each primer pair using the NEB Annealing Temperature Calculator (New England Biolabs, 2025) and further optimised through running a gradient PCR. Gradient PCRs were run with annealing temperature set at least ±2°C from the recommended value to identify conditions that produced the strongest amplification, as assessed by gel electrophoresis.

`

2.4.2 Quantitative PCR (qPCR)

All quantitative PCR (qPCR) assays were performed using the Rotor-Gene 6000 system (Qiagen). Each reaction was prepared using TB Green Premix Ex Taq (Takara) in a final volume of 20 µL, consisting of: 10 µL 2× TB Green Premix Ex Taq, 1 µL of 10 µM forward primer, 1 µL of 10 µM reverse primer, 3 µL of nuclease-free water and 5 µL template (either gDNA or cDNA). PCR cycling conditions varied depending on the specific application but typically consisted of an initial denaturation step of 98°C for 30 seconds, followed by 35 cycles of 95°C for 5 seconds, 60°C for 15 seconds, 72°C for 15 seconds (first acquisition; cycling A) and 84°C for 15 seconds (second acquisition; cycling B). Reactions concluded with a final ramp from 60-99°C for 5 seconds to generate a melt curve.

2.4.3 Digital Droplet PCR (ddPCR)

All digital droplet PCR (ddPCR) procedures were performed using the Bio-Rad QX200 Digital Droplet PCR system. Each reaction mix was prepared using 1.1 µL template DNA, 11 µL of 2× EvaGreen Supermix, 2.2 µL 10 µM forward primer, 2.2 µL 10 µM reverse primer and 5.5 µL of water. Droplet generation was carried out using 20 µL of this reaction mix as per manufacturer's instructions employing the Bio-Rad QX200TM droplet generator, DG8 cartridge holders, DG8 cartridges, DG8 gaskets and QX200 Droplet Generation Oil for EvaGreen. Generated droplets were transferred into semi-skirted 96-well PCR plates and seal with pierceable heat-seal foil using PX1 PCR plate sealer at 180°C. Amplification of DNA targets was performed in a Bio-Rad C1000 Touch Thermal Cycler under the following conditions: initial denaturation at 95°C for 5 minutes, 40 cycles of 95°C for 30 seconds then 60°C for 1 minute, 4°C for 5 minutes, 90°C for 5 minutes, then a final hold at 12°C. Ramping time for each step was set at 2°C/second, with the final ramp to hold temperature at 1°C/second. Following thermal cycling, droplets were analysed using the QX200 Droplet Reader according

`

to the manufacturer's protocol with ddPCR Droplet Reader Oil. Plate details, run configuration and data analysis were performed using Bio-Rad Quantasoft software (version 1.74.0917). Samples that were oversaturated (copies/µL exceeding 5000) or had insufficient accept droplets ($\leq$12,000) were excluded from analysis.

2.4.4 Digital PCR (dPCR)

All digital PCR (dPCR) procedures were performed using the QIAcuity instrument (Qiagen) using EvaGreen Supermix and were run in accordance with manufacturer's instructions using 8.6K nanoplates. The dPCR reaction mix consisted of 4 µL EvaGreen Supermix, 0.5 µL of 10 µM forward primer stock, 0.5 µL of 10µM reverse primer stock, 6 µL water and 1 µL sample. Solutions were prepared in strip tubes and then transfer into the nanoplate. The dPCR settings were as follows: initial denaturation at 95°C for 2 minutes, followed by 40 cycles of 95°C for 15 seconds, 60°C for 15 seconds and 72°C for 15 seconds. A final extension step was performed at 40°C for 5 minutes. Imaging was carried out with a green channel exposure time of 300 ms at gain 6. Samples exhibiting oversaturation (>2000 copies/µL) or had insufficient accepted partitions (<6000) were excluded from analysis.

2.5 rAAV vector production

2.5.1 Production of rAAV by calcium phosphate triple transfections

Recombinant AAV vectors (rAAV) were produced via triple plasmid transfection of HEK293 cells using calcium phosphate transfection. HEK293 cells were seeded at a density (9.4×10^6) in 15 cm cell culture dish (Corning) in 20 mL growth media such that they reached 80% confluency the following day.

`

On the day of transfection, cell medium was aspirated and replaced with transfection media to encourage cell division. Transfections were performed two hours later. Each dish received 5 μg of vector plasmid, 9 μg capsid plasmid, and 18 μg adenoviral helper plasmid. Transfection solutions A and B were prepared in each ratio to a total volume of 2.5 mL per dish. Solution A contained 0.25 M $CaCl_2$ and 32 μg total plasmid DNA, made up to a total volume of 2.5 mL in 10% TE. Solution B contained 1.5 mM NaH_2PO_4 in 2×HBS. Solution B was aliquoted into 2.5 mL volumes and 2.5 mL of Solution A was added dropwise while agitating to form a fine precipitate. Transfection mix was incubated for 15-30 minutes at RT then 2.5 mL was added to each culture dish.

The following day, cell medium was aspirated and replaced with maintenance media. Cells were harvested 48 hours later (three days post transfection) using a cell lifter (Corning) and transferred into sterile 500 mL bottles (Corning). Cells were pelleted by centrifugation at 15000 rpm for 5 minutes. The supernatant was transferred into fresh 500 mL bottles (Corning), and cells resuspended in benzonase buffer (~0.3 mL per dish). Both the media and cells fractions were stored at –80°C until purification.

2.5.2 rAAV purification through CsCl gradients

Recombinant AAV virions were purified using a caesium chloride (CsCl) gradient purification method, with all relevant steps conducted in a biosafety hood.

Frozen cells and media were thawed in a 37°C water bath and processed separately.

Cells underwent three freeze-thaw cycles to induce lysis followed by treatment with benzonase endonuclease ($\geq$125 units/mL). The solution was incubated at 37°C for 1 hour and mixed using intermittent vortexing. Lysates were centrifuged at 3,700×*g* for 30 minutes at 4°C to pellet cell debris and the supernatant was transferred into a fresh 50 mL falcon and

`

1/39th volume 1M $CaCl_2$ added. This solution was incubated on ice for 1 hour before centrifugation at 3,700×*g* for 30 minutes at 4°C. The supernatant was again transferred into a fresh 50 mL Falcon tube. Both the clarified cell lysate and thawed media were supplemented with 1/4th volume 40% (w/v) PEG 8000 2.5 M $NaCl_2$, mixed thoroughly and incubated on ice overnight.

The following day, suspensions from the cell and media fractions were centrifuged at 3,700×g for 30 minutes at 4°C. The supernatant was discarded, and the resulting white precipitate containing the AAV vector, resuspended in resuspension buffer, being careful to break up any clumps as much as possible. Resuspended solutions were transferred into 50mL falcon tubes and placed on a tube rotator overnight at 4°C. Typically, media pellets were combined in two separate suspensions of 20 mL and cell pellets were combined into one total suspension of 20 mL.

The first caesium chloride gradient was prepared in 25 × 89mm ultracentrifuge tubes (Beckman Coulter) by layering 5mL of 1.5 g/mL CsCl under 12 mL of 1.3 g/mL CsCl to form a clear density interface. The viral suspension (20 mL) was carefully layered on top. Tubes were balanced using resuspension buffer, loaded into a SW32 Ti rotor and centrifuged for 24 hours at 106,800×*g* at 20°C.

After centrifugation, vector bands were extracted using a 10 mL syringe with an 18G needle. The needle was inserted at the junction of the rounded bottom and side wall of the tube, and 5-6 mL of clear CsCl band (between the 1.5g/mL caesium cushion and the protein debris layer) was collected. The vector solution was transferred to a 14 x 89 mm Ultra-Clear ultracentrifuge

`

tube (Beckman Coulter) and adjusted to a final density of 1.37 g/mL. Tubes were filled to the top, balanced and loaded into a SW41 Ti Rotor and spun at 247,600×*g* at 20°C for 48-60 hours. Tubes were removed from ultracentrifuge, sealed with parafilm and placed in a retort stand. An 18G needle was used to puncture the bottom of the near the curvature and 1mL fractions were collected dropwise. Typically, two AAV bands were visible, the lower representing full AAVs and the upper empty AAV capsids. All fractions were stored at 4°C while vector content was quantified using qPCR analysis (Section 2.5.3).

Fractions with the highest vector concentration (peak and one fraction above and below) were pooled and loaded into a pre-wettened Slide-A-Lyzer Gamma Irradiated Dialysis Cassette (Thermo Scientific) using an 18G needle and syringe. Cassettes were placed in 4 L dialysis solution spinning overnight at 4°C on a magnetic stirrer. Cassettes were changed into fresh dialysis solution over the following day before transfer into a third and final dialysis solution containing 5% (v/v) glycerol for overnight incubation.

Following dialysis, virus was removed from the cassette, transferred into a Vivaspin 20 100,000 MWCO (Satorius) and concentrated by centrifugation at 3,700×*g* to a final volume between 200-1000 µL. Virus was then stored at 4°C for later titration via ddPCR and experimental use.

2.5.3 Quantification of vector particles in viral fractions by qPCR

AAV viral fractions from Section 2.5.2 were quantified by qPCR against known standards to estimate the number of viral genomes present in each fraction. Standard qPCR protocol as described in Section 2.4.2 as followed.

`

Vector fractions were thoroughly mixed by pipetting, and 2 µL of each was diluted 1:500 in nuclease-free water to serve as the template DNA. Standards were prepared from the packaged viral plasmid at concentrations ranging from 1×10^3 to 1×10^8 per 5 µL, using serial 1:10 dilutions. Reactions were performed using bovine growth hormone poly adenylation signal (BGHpA) primers (Table 8). Each standard was typically run in duplicated to improve the accuracy of the standard curve calculation.

2.5.4 Quantification of purified vector stocks by ddPCR

AAV vector titre was determined using Bio-Rad's ddPCR system as described above. Vector samples were denatured by incubation at 99°C for 10 minutes in ddPCR alkaline digestion buffer, using 2 µL vector to 49 µL buffer. Following thermal denaturation, samples were rapidly cooled to 4°C for 1 minute on ice. Equal volume of cold neutralisation buffer was then added to neutralise the digestion buffer. Serial dilutions were subsequently performed to obtained final vector concentrations of 1:1,000, 000 and 1:4,000,000 relative to the initial vector stock. These dilutions served as the template DNA for ddPCR protocol (Section 2.4.3). BGHpA primers (Table 8) were used for viral titre as every vector contained a bovine growth hormone polyadenylation sequence.

2.6 Constructs of rAAV plasmids and *in vitro* testing

2.6.1 Construction of ssAAV-CRISPR/Cas9 vectors with sgRNAs

Candidate guides oligos were designed as described in detail in Section 3.2.3. Two different SaCas9 plasmids were used for *in vitro* and *in vivo* studies respectively. The plasmid backbone pX601-AAV-CMV::NLS-SaCas9-NLS-3xHA-bGHpA;U6::BsaI-sgRNA (211) was used for *in vitro* guide testing in murine cell lines NIH3T3 and B35 (Table 2.4), with SaCas9 expression under a ubiquitous promoter.

`

The plasmid backbone pX602-AAV-TBG::NLS-SaCas9-NLSHA-OLLAS-bGHpA;U6::BsaI-sgRNA (211) was used for *in vivo* testing in PKU mice. In this plasmid, the expression of the SaCas9 transcript is driven from the liver-specific human thyroxine binding globulin promoter (TBG) promoter to limit extra-hepatic expression.

Guide RNA sequences were orders as plus (+) and minus (-) strand oligonucleotides from Sigma (Table 2.10), flanked by compatible cohesive ends. Plus (+) and minus (-) strand oligonucleotides were phosphorylated and annealed in a thermocycler. The annealed and phosphorylated gRNA oligonucleotides were then ligated into the required plasmid backbone that has been digested to completion with BsaI restriction enzyme (New England Biolabs). Plasmids were Sanger sequenced to identify clones that contained the correctly inserted gRNA sequence.

Table 2.10: Top and bottom oligonucleotides coding each gRNA sequence.

gRNA	Strand	Sequence 5'-3'
1	+	CACCGTAGCACATGCCGATATGCACC
	-	AAACGGTGCATATCGGCATGTGCTAC
2	+	CACCGCCGGGATTGGCTAGGTAAACC
	-	AAACGGTTTACCTAGCCAATCCCGGC
3	+	CACCGGGGCCATGTTACATCATAAAT
	-	AAACATTTATGATGTAACATGGCCCC
4	+	CACCGCAGGCTTTCTTACGTGTTAGC
	-	AAACGCTAACACGTAAGAAAGCCTGC

Coloured = BsaI restriction enzyme cleavage sites compatible cohesive ends

2.6.2 *In vitro* guide screening

The pX601-AAV-CMV + gRNA plasmids were transfected into either NIH3T3 or B35 cells using either calcium phosphate transfection or lipofectamine. Cells were seeded for following day 80% confluence: 8 × 10^5 cells/well in a 6 well plate in 2 mL cell media. The following day, media was changed to Transfection media at least 1 hour prior to using one of the transfection approaches. A range of plasmid DNA concentrations were attempted to try to optimise either approach.

For calcium phosphate transfections, as previously described in rAAV manufacturing (Section 2.5.1), solutions A and B were prepared in a total volume of 150 µL of each solution per well. Solution A contained 0.25M CaCl2, plasmid DNA, made up the total volume with 10% (v/v) TE Buffer (Table 2). Solution B contained 15mM Na2HPO4 in 2 x HBS (Table 2). Solution A was added dropwise to solution B while stirring constantly to form a fine precipitate. The resulting solution was incubated for 15-30 minutes are RT and then 300 µL was added to each well of cells.

Lipofectamine (2000) transfections were performed as per manufacturer's instructions. For both approaches, a CMV-GFP plasmid was used as a transfection control as standard in our laboratory.

2.6.3 Construction of scAAV HITI donor vectors

Coding sequences in all donor plasmids were codon-optimised for human expression using an *in silico* tool (GenScript, 2025) and were designed to exclude PacI and SpeI restriction enzyme sites. Input sequences were either the full human *PAH* cDNA sequence, human *PAH* exons 2-13 cDNA or full EGFP cDNA. Codons were optimised to improve human expression as previous studies had shown human PAH functional in the mouse and for better clinical translation. Sequences were codon optimised to enhance protein translation efficiency by matching codon usage patterns of the host organism (212). Human PAH sequence was used for more direct clinical translatability and has been shown in many previous studies to be able to correct phenotype in mice (141). Overview of each donor approach shown in Figure 3.2.

`

Custom donor sequences (Table 7.1, Appendices) were synthesised by GenScript (Nanjing, China) provided in the pUC57-Kan backbone and subsequently cloned into a plasmid backbone containing AAV ITRs (pAAV2-LPhOTC BB2; Appendices 6.1.1) as described briefly below.

All donor sequences were designed to be flanked with PacI and SpeI sites for easy cloning. The PAH-Splice donor sequence, the PAH-2PA donor sequence, the GFP-Splice donor sequence and the GFP-2PA donor sequence (Table 6.1, Appendices) were cut from the backbone using PacI and SpeI double restriction digest and visualised by gel electrophoresis to confirm expected band size.

Excised donor sequences were ligated into the pAAV2-LPhOTX BB2 backbone that had been linearised with PacI and SpeI to generate compatible termini for cloning (A/Prof Samantha Ginn). Resulting plasmids were Sanger sequences to confirm correction insertion. Final plasmids stocks were digested with AhdI, AvaI, MscI, and XmaI restriction enzymes and the size of the resulting bands were analysed by agarose gel electrophoresis to confirm AAV ITR integrity.

2.7 Murine model of PKU

All animal experiments were approved by the CMRI Animal Ethics Committee (C400, approved 14th August 2023). Mice were housed in open air or individually ventilated cages under a standard 12-hour light/dark cycle. Mice were given chow and water *ad libitum*. Two chows were used in this project: standard (20% (w/w) protein, SF00-100, Speciality Feeds) or high protein (61% (w/w) protein, SF18-142, Speciality Feeds). Methods for diet-based phenylalanine elevation, including high protein chow and phenylalanine-supplemented water, were evaluated as potential methods for dietary phenylalanine elevation. Animals were

`

monitored daily by CMRI BioResources staff. Health monitoring was performed quarterly by testing common pathogens.

Heterozygous B6.BTBR-*Pah*^{enu2}/MalnJ breeding pairs were obtained from The Jackson Laboratory (029218; Bar Harbor, USA). The Pah^{enu2} allele carries a missense mutation in exon 7 of *Pah* (c.835T>C), originally generated by chemical mutagenesis in BTBR mice (116). Although this specific mutation has not been documented in patients, homozygous Pah$^{enu2/enu2}$ mice are known to display severe hyperphenylalaninemia and neuropathology consistent with human disease.

Because the BTBR background is associated with innate neurological phenotypes, the line had previously been backcrossed for multiple generations onto a C57BL/6 by Dr Michael Allen (University of North Texas Health Science Centre). The model is referred to hereafter as *Pah*^{enu2}.

Heterozygous mice were initially bred to produce homozygous *Pah*^{enu2}/^{enu2} mice. Once the colony was established, heterozygous females were crossed with homozygous mutant male mice for ongoing colony maintenance. This breeding strategy maximised the number of homozygous mutant mice produced while avoiding maternal PKU effects, as homozygous females have reduced litter sizes and offspring viability.

2.7.1 Genotyping

Genotyping was performed on ear notch samples, obtained as a bioproduct of animal numbering, using hot sodium hydroxide for crude DNA extraction, also known as HotSHOT DNA extraction (204). Briefly, 75 µL of HotSHOT alkaline digestion buffer was added to

`

each ear notch and incubated at 95°C for 1 hour. The reaction was cooled to 4°C before adding 75 µL neutralisation buffer. Samples were used immediately, with minimal shelf life at 4°C.

Probe qPCR was performed using crude DNA extract. Protocol is publicly available on JAX (The Jackson Laboratory, 2025). Probe FAST qPCR mastermix (KAPA BioSystems) was used for each PCR reaction, prepared as 5 µL 4× Probe FAST qPCR mastermix, 1 µL of a 20× primer/probe mix, 9 µL nuclease-free water, and 5 µL of crude DNA extraction. 20× primer/probe mix had a final concentration of 8 µM CL013F primer, 8 µM CL014R primer (Table 8), 3 µM WT probe and 3 µM Mutant probe (Table 9). A 500 µL mix was prepared and aliquoted at the beginning of the project to prevent repeated cycles of freeze/thaw. Qiagen Rotorgene cycler was used to run qPCR. PCR conditions were set as an initial denaturation step of 95°C for 2 minutes, followed by 40 cycles of 95°C for 5 seconds, 60°C for 30 seconds. Fluorescence (cycling A) was acquired after each cycle in both the green and yellow channel. Scatter graph analysis was performed to compare relative Cycling A. Green vs Cycling A. Yellow intensities to clearly separate three genotypes possible.

2.7.2 Blood sampling

Blood was taken via submandibular puncture fortnightly and prior to sacrifice in accordance with animal ethics approval. Mice fasted for 4 hours beforehand to control for dietary phenylalanine intake. Mice were securely restrained, and a 25-gauge needle was used to puncture the right facial vein. The resulting blood droplet was blotted directly onto filter paper, identical to that used for newborn screening at the Children's Hospital in Westmead. Blood samples were allowed to dry overnight, after which a 3 mm disc was punched from each spot

`

and placed in 98-well plate for blood spot analysis. Following sampling, blood spot cards were store in room temperature in the dark.

2.7.3 Blood phenylalanine and tyrosine quantification

Blood phenylalanine and tyrosine were analysed using liquid chromatography-tandem mass spectrometry (LC-MS/MS) by hospital scientists at Newborn Screening (NBS) at the Children's Hospital, Westmead. The NeoBase 2 Non-derivatised MSMS Kit (Revvity) was used following manufacture's guidelines. Blood analytes were extracted from a 3 mm dried blood spot punch with 125 µL extraction working solution (EWS) in a microplate and shaken for 30 minutes at 45° C (650-750 rpm). EWS contains an internal standard containing stable isotope-labelled phenylalanine ($^{13}C_6$-Phe) and isotope-labelled tyrosine ($^{13}C_6$-Tyr). To a fresh microplate, 90 µL of extraction solution and 10 µL of extract was added using the Andrews pipetting robot (Waters Corporation). After sealing the plate, it was loaded into the autosampler for analysis. Analytes were quantified using the MassLynx 4.2 Software (Waters Corporation). Quantification is performed by correction using internal standards. Data is compared to reference range.

2.7.4 Delivery of rAAV vectors to mice

Recombinant AAV vectors were diluted in sterile sodium chloride (0.9% w/v) to a total of 200 µL and injected intravenously (i.v.) via the tail vein. Mice were warmed under a heat lamp for a maximum of ten minutes to dilate the tail vein prior to injection. An orange light source (Veinlite) was used as a visual aid, and mice were restrained in a plastic rodent restrainer. Mice were monitored following injection for any adverse events. For *in vivo* guide testing, mice were injected with 5×10^{11} of SaCas9 vectors with respective candidate guides. For donor testing, mice were injected with 1×10^{12} SaCas9 vectors alongside 5×10^{11} donor vector, or 5×10^{11} donor vector alone. Vector doses were selected based on previous laboratory data identifying the dose and dose ratios that achieve optimal targeted integration using this approach with similar editing reagents (213).

2.7.5 High protein diet challenge protocol

Mice undertaking a high protein diet had first blood taken for baseline phenylalanine and tyrosine analysis via submandibular puncture and placed in new cages free of any other dietary sources. They were fed a 61% (w/w) protein diet until sacrifice.

2.7.6 Standard harvest protocol

All animals were culled humanely either via cervical dislocation or isoflurane overdose when performing liver perfusions. Upon sacrifice, liver was harvested. Small liver blocks were taken for fixation in 4% (w/v) PFA overnight at 4°C and the remainder of the liver was homogenised and snap frozen at liquid nitrogen for DNA, RNA and protein extraction. Following fixation, liver blocks were cryoprotected by sequential incubations in a sucrose gradient. PFA was removed and replaced with 10% (w/v) sucrose in PBS and samples were incubated overnight on a rotator at 4°C. This process was repeated with 20% (w/v) and 30% (w/v) sucrose in PBS. Excess sucrose was blotted off using paper towel and samples were embedded in Tissue-Tek Optimal Cutting Temperature Compound (O.C.T) (Sakura) within plastic moulds. Embedded moulds were snap-frozen by immersion in 2-methylbutane cooled with liquid nitrogen and stored at –80°C.

2.7.7 Isolating of murine hepatocytes by perfusion

Murine hepatocytes from PKU mice were isolated either manually via collagenase perfusion or using the automated gentleMACS perfusion system (Miltenyi Biotec).

`

2.7.7.1 Manual collagenase perfusions

Mice were euthanised by isoflurane overdose to minimise damage to circulation used for the perfusion. The abdominal cavity was exposed and the left lateral liver lobe ligated at the base using cotton thread. Perfusion machine as set up the administer solutions through tubing connected to a 25G needle. This needle was carefully inserted into inferior vena cava and secured with a surgical clamp. The hepatic portal vein was cut to prevent excessive pressure from causing damage to blood vessels. After confirming the system was set up correctly, the lobe was excised for molecular analysis, and perfusion was initiated. A series of warmed solutions (37°C) were passed through the mouse circulatory system at a flow rate of 12. These solutions were: 25 mL solution 1 (HBSS -/-; to flush out red blood cells), 25 mL solution 2 (0.5 mM EDTA; to dissolve cell-cell adhesion), 25 mL solution 3 (HBSS (-/-); to remove EDTA) and solution 4 (5 mM $CaCl_2$ HBSS containing 0.05% (w/v) collagenase and 0.01% (w/v) DNaseI). Successful perfusion was indicated by a colour change in the liver to pale yellow. The liver was then carefully removed from the abdominal cavity and transferred into a 10 cm petri dish containing ice cold 10% DMEM, ensuring that the gallbladder was not included. Hepatocytes were carefully dissociated by scraping liver tissue with the blunt end of a razor blade. The resulting solution was transferred into a 50 mL falcon tube and placed on ice until further processing.

2.7.7.2 Automated perfusions using gentleMACS perfuser

For automated perfusions, the gentleMACS system was used according to the manufacturer's instructions using the mouse and rat liver perfusion kit (Miltenyi Biotec). Briefly, mice were euthanised by cervical dislocation, and the largest liver lobe (left lateral lobe) was dissected, taking care not to damage the capsule (outer membrane of the liver). The liver lobe was stored on ice either in PBS for short term use ($\leq$10 minutes) or in Tissue Storage Solution for longer-term storage (≤20 minutes). The liver lobe was placed flat side down in the centre of the perfuser grid and secured with a clamp. The assembled perfuser was then placed in the GentleMACS instrument, with a heating unit positioned over it. The perfusion program consisted of four short washed (30 seconds each) followed by one long wash (12 minutes) in pre-digestion buffer, a 30 second wash with equilibration buffer and 10 minutes in enzymatic buffer. All solutions were warmed at 37°C prior to use. The contents of perfuser were transferred in a mixing tube (C tube), where the hepatocytes were released into the solution by mixing for 5 minutes. The cell suspension was filtered through a 100 µm strained and 6 mL cold 10% DMEM added. Cells were placed on ice for subsequent processing.

2.7.7.3 Processing perfused murine hepatocytes

Regardless of perfusion method, the hepatocyte solution was filtered through a 100 µm strainer to remove any large tissue fragments, and the solution centrifuged at 400×*g* for 3 minutes at 4°C. The supernatant was discarded and pelleted cells were resuspended in 21 mL 10% DMEM with 9 mL of isotonic Percoll. The solution was mixed thoroughly through inversion and centrifugated again at 860 ×*g* for 10 minutes at 20°C. The supernatant was discarded, and the cell pellet was resuspended in 10 mL 10% DMEM.

An aliquot of cells was diluted 1:10 in Trypan Blue, incubated for 5 minutes and then counted using hemacytometer as described. Cells were centrifuged at 400 ×*g* for 3 minutes at 4°C, supernatant removed and resuspended in cold CS10 at a density approximately 5×10^6 cells/mL. Finally, cells were cryopreserved as described (Section 2.2.3).

`

2.8 Liver tissue processing

2.8.1 Phenol-Chloroform DNA extraction from liver tissue

Genomic DNA was extracted from snap-frozen murine liver tissue. Liver samples were homogenised with a rotary pestle and electric hand-held homogeniser with 300 µL liver lysis buffer (Table 2) for at least 30 seconds. An additional 300 µL liver lysis buffer and 5 µL RNAse A (20 mg/mL) were added and incubated for one hour at 37°C. 5uL Proteinase K (20 mg/mL) was added and samples were incubated overnight at 56°C.

The following day 600 µL phenol:chloroform:isoamyl alcohol (25:24:1 ratio) was added to each sample, mixed by inversion 5 minutes and centrifuged for 10 minutes at 14000×g. The aqueous phase was then collected and the preceding step repeated once more. To remove residual phenol, an equal volume of chloroform:isoamyl alcohol (24:1 ratio) added to the aqueous phase, mixed by inversion for 5 minutes and centrifuged for 10 minutes at 14000×g. A volume of 400 µL from the aqueous layer was taken into a fresh Eppendorf tube and precipitated using 80 µL 10 M ammonium acetate (0.25× volume) and 1000 µL cold 100% ethanol (2.5×volume). Samples were vortexed thoroughly and incubated at –20°C for a minimum 30 minutes prior to 30 minutes centrifugation at 14000×g at 4°C to pellet DNA. The supernatant was discarded, and DNA pellets were washed three times with 70% ethanol prior to air drying to remove residual ethanol and resuspension in 10% TE buffer. DNA was left at room temperature overnight to fully dissolve. Resuspended DNA was stored at –20°C until further use.

`

2.8.2 RNA extraction and cDNA synthesis

2.8.2.1 RNA extraction

Total RNA from snap-frozen liver tissue (50-100 mg) using a modified TRIzol-based protocol in combination with the PureLink RNA Mini Kit (Invitrogen). Tissue was homogenised in 200 µL TRIzol using a rotary pestle and electric homogenizer (30 seconds), followed by the addition of 800 µL TRIzol and incubation at room temperature for 5 minutes. Phase separation was achieved by adding 200 µL chloroform, shaking vigorously (15 seconds), and incubating for 3 minutes before centrifugation (12,000×g, 15 seconds, 4°C). The colourless aqueous phase (600 µL) was transferred into a fresh tube with equal volume of 70% ethanol, vortexed and loaded into a PureLink spin column in 700 µL aliquots after entire sample loaded via centrifugation (12,000×g, 15 seconds). The column was washed with 250 µL Wash buffer I and centrifugation at 12,000 ×g for 15 seconds. On-column DNase treatment (80 µL, PureLink DNase Kit) was performed according to the manufacturer's instructions and incubated for 15 minutes at room temperature. A second wash with 350 µL Wash buffer I was followed by two washes of 500 µL of wash buffer II via centrifugation (12,000×g for 15 seconds). A final dry spin (12,000×g, 15 seconds) was performed before elution with 90 µL RNase free water (1 minute incubation, then centrifugation at 12,000×g for 15 seconds). RNA yield and purity were assessed using NanoDrop spectrophotometer, and samples were aliquoted and stored at -80°C.

2.8.2.2 Synthesis of complementary DNA (cDNA) from total RNA

RNA samples were used to synthesis cDNA using SuperScript IV First-Strand Synthesis System (Thermo Scientific). Prior to cDNA synthesis, 2 µg of RNA was incubated with DNase I for 30 minutes at 37°C to remove residual genomic DNA. DNase-treated RNA was then divided into two aliquots to generated paired reactions with and without reverse transcriptase enzyme. Approximately 1 µg of template RNA was used per reaction, and cDNA synthesis was performed using the Oligo(dT) method according to manufacturer's instructions. Resultant cDNA was stored at –20°C until use.

2.8.3 Liver protein extraction and quantification

Liver tissue (50 mg) was homogenised in 200 µL cold Thöny lab protein extraction buffer (Table 2) using rotary pestle and electric homogeniser for 10 seconds. 500 µL Thöny lab protein extraction buffer was add before centrifugation at 13,000 ×g for 30 minutes at 4°C. A small aliquot of protein lysate was taken for quantification and the rest stored on dry ice until use.

Protein lysates yield was determined using the DC Protein Assay Kit (Bio-Rad) utilising a colourimetric assay based on Bradford protein method. Kit was used as per manufacturer's instructions. Briefly, protein lysate samples were diluted serially, minimum of two dilutions (neat; 1 in 10; 1 in 40). Protein standards were included with each assay (range of 0-7 µg/rxn), and samples were run in triplicate on a 96-well plate. Protein Assay Kit was used as per manufacturer's instructions. Plates were read at 600 nm with the GlowMax Microplate Reader (Promega) at room temperature. Average values were taken across triplicate standards, and a standard curve was used to estimate the concentrations of the samples accounting for dilution factors.

2.8.4 Sectioning of cryopreserved tissue

Cryo-preserved liver tissue blocks were mounted onto cutting blocks and section at 5 µM sections using Leica Cryostat onto Menzel-Glaser Superfrost Plus Microscope Slides (Thermo Scientific). Tissue blocks were placed at –20°C for at least 1 hour prior to sectioning and transported on dry-ice at the conclusion of sectioning. Slides with cut sections were stored at 4°C short-term before staining. Slides were typically used within days of being cut.

2.9 Molecular analyses of murine liver samples

2.9.1 Quantification of liver tissue vector copy number using dPCR and ddPCR

Vector copy number was undertaken using genomic DNA extracted as per methodology for Phenol-Chloroform DNA extraction. DNA yield was determined using Qubit BR dsDNA kit (Thermo Fisher Scientific) and 24 ng total DNA was digested enzymatically using HindIII (New England Biolabs) as per manufacturer's instructions for 2 hours at 37°C. DNA digest was diluted 1 in 20 (1.2 ng DNA) and amplified using primers (Table 8) 330_F and 331_R for Cas9 vectors, GFP_F and GFP_R for GFP vectors and CL300F and CL301R for PAH vectors either using described ddPCR or dPCR methodologies (Section 2.4.3 and Section 4.4.4). Signals were normalised for each sample with a parallel PCR targeting murine albumin using primers mAlb_F and mAlb_R (Table 8) using undiluted DNA template (24 ng), accounting for amount of input DNA. dPCR analysis was adopted part way through due to consistent faults to ddPCR machines and provides equivalent results.

2.9.2 Genomic DNA junction analysis using dPCR

Junction events were quantified using the Qiagen digital PCR (dPCR) system, following the general protocol described in Section 2.9.2. EvaGreen mastermix was used according to the manufacturer's instructions. Genomic DNA (880 ng) was digested enzymatically using HindIII (New England Biolabs) in a total volume of 50 µL for 2 hours at 37°C. Template DNA concentration of 132 ng.

To assess the orientation of the insert at both the 5’ and 3’ junctions, primers (Table 8) were designed such that one primer targeted endogenous genomic DNA and the other targeted the inserted fragment. This allowed for the detection of four configurations: correct orientation at the 5’ end, correct orientation at the 3’ end, reverse orientation at the 5’ end and reverse orientation at the 3’ end. Results were normalised to murine albumin. Primers spanning the cut site were used as a reference for genomic integrity.

2.9.3 Transcript analysis by dPCR

Transcripts were quantified using the Qiagen digital PCR (dPCR) system, following the general protocol described in Section 2.9.2. EvaGreen mastermix was used according to the manufacturer’s instructions. To assess the hybrid transcript, primers (Table 8) were designed such that one primer targeted endogenous exon 1 and the other targeted integrated sequence. Results were normalised to *Hprt*. Primers specific to endogenous murine *Pah*, which uses the same promoter, were used to calculate total proportion of transcripts per locus. Template cDNA was diluted 1:50 prior to assay.

2.9.4 PAH *in vitro* activity assay

For each reaction, 25 μg of protein was combined with 50 μL of 2× substrate solution (final assay concentration:150 mM potassium chloride, 400 μg/mL catalase, 300 mM ammonium iron (II) sulfate, 2 mM dithiothreitol, 200 mM phenylalanine-d5, and 200 μM tetrahydrobiopterin in aqueous solution) and 50 μL of 2× enzyme solution (final assay concentration: 150 mM potassium chloride, 400 μg/mL catalase, 300 mM ammonium iron (II) sulfate, 2 mM dithiothreitol, and 225 μg/mL liver extract) in a 96-well microplate. Reactions were incubated at room temperature for the specified time to permit conversion of phenylalanine to tyrosine, then quenched with 10 μL of 10% formic acid containing 5 μM tyrosine-d4.

2.9.4.1 Preparation for MS

Samples were frozen until processing. Prior to extraction, samples were thawed and transferred to 1.5 mL tubes. Methanol (500 μL) was added to each tube, followed by vortexing for 1 minute and cooling at 4°C for 30 minutes. Following a second 1-minute vortex, tubes were centrifuged at 18,000 ×g for 15 minutes. The supernatant was transferred to a new 1.5 mL tube and dried in a vacuum concentrator. Dried samples were reconstituted in 50 μL 0.1% formic acid 20 mM ammonium formate solution, then centrifuged at 12,500 ×g for 1 minute in a microfuge. A 5 μL aliquot of the resulting supernatant was loaded into a 96-well plate for LC-MS/MS. Preparation for mass spectrometry was performed by Dr Mark Graham.

2.9.4.2 Mass spectrometry

The LC-MS/MS was performed using a Dionex UltiMate 3000 HPLC system and Q Exactive Plus hybrid quadrupole-orbitrap mass spectrometer (Thermo Fisher Scientific). Each sample was loaded onto a Luna 50 × 2 mm, 2.5 um particle C18(2)-HST column (Phenomenex, PN 00B-4446-B0), with a 2.1 mm guard column (Security Guard Ultra C18, Phenomenex, PN AJ0-8782). The column was heated to 50°C using the column oven module. The sample was injected as 1 μl and loaded onto the column in 99% phase A (solution of 0.1% formic acid, 20 mM ammonium formate in water) and 1% phase B (solution of 0.1% formic acid, 20 mM ammonium formate in 90% acetonitrile). The flow rate was 200 μL/min. The sample loop was in-line for 1 minute with 99% A and then bypassed. The gradient was from 1% B to 10% B in 5.5 minutes, to 90% buffer B in 1 minute, and held at 90% buffer B for 1 minute, to 1%

`

buffer B in 1 minute and held for 2.5 minutes to re-equilibrate the column. The entire HPLC run was 10 minutes. Each sample was run once with a blank run in between. The standards were run twice. The HESI-II ionisation probe was used to allow electrospray to operate at 4 kV. The S lens radio frequency level was 70 and capillary temperature was 300 °C. The sheath gas flow was set at 30 L/min. The auxiliary gas flow was 10 L/min. The sweep gas flow was 2 L/min. The auxiliary gas temperature was 300 °C. Both MS and MS/MS spectra were acquired for 10 minutes. The positive MS scan was from 100-400 m/z for 200 ms with an automatic gain control target of 3,000,000 counts and resolution of 70,000.

The MS/MS spectrum was targeted at 931.5896 m/z of sirolimus with an ammonium ion. Each targeted MS/MS scan was 150 ms, with an automatic gain control target of 500,000 counts, resolution of 35,000 and isolation window of 0.4 m/z. The first mass was dynamic, and the normalised collision energy was 30. The inclusion list contained D5-Phe at 171.1776 m/z and D4-Tyr at 186.1063 m/z. The raw LC-MS/MS data was processed with Skyline 25.1.0.142. For D5-Phe, the D5-immoium ion at 125.1122, and for D4-Tyr the D4-immnium ion at 140.1008, were the transition used for quantification. The intensity data for these transitions was exported to Microsoft Excel. Dr Mark Graham and Dr Md Musfizur Hassan performed mass spectrometry.

2.9.5 Flow cytometric analysis of GFP expression in hepatocytes

Murine hepatocytes perfused from PKU mice were recovered from cryopreservation as described in Section 2.2.3. After recovery, cells were centrifuged at 40×g for 3 minutes, supernatant removed and the pellet was resuspended in 1 mL FLOW buffer (Table 2). Hepatocytes were transferred into round-bottom tubes with 35 μM cell strainer caps (Corning) and stained using 5 μL of DAPI (20 mg/μL) on ice for 5 minutes in the dark.

`

Samples were analysed on an LSRFortessa (BD Biosciences) equipped with a 405 nm violet laser for DAPI detection and a 488 nm blue laser for eGFP detection. DAPI fluorescence was detected using a 450/50 bandpass filter, and eGFP was detected using a 530/30 filter. Instrument settings were optimised using untreated controls. Flow cytometric acquisition and subsequent data analysis in FlowJo (BD Biosciences) were performed by Cindy Zhu. Initial gating was conducted on forward scatter (FSC-A) versus side scatter (SSC-A) to identify hepatocyte-sized events and exclude debris. Doublets were removed by gating on SSC height (SSC-H) versus SSC-A. Live cells were identified by the exclusion of DAPI-positive events. GFP expression was assess in live singlets by plotting FSC-A versus GFP (530/30). The GFP-positive gate was defined based on the fluorescence distribution of an untreated mouse sample (GFP-negative). A minimum of 10,000 live single-cell events were collected per sample. Quantitative outputs were percentage GFP-positive live single hepatocytes and the mean fluorescent intensity (MFI) of GFP positive cells. The gating strategy is shown in the Appendices (Section 6.1.4).

2.9.6 Immunofluorescent co-staining of GFP and GLUL in frozen liver sections

Sections were washed with PBS for 5 minutes, then fixed in ice-cold methanol for 20 minutes at 20°C. Following fixation, sections were washed for 10 minutes in 1×PBS, permeabilised in 0.1% (v/v) Triton X-100 in PBS for 10 minutes at room temperature and washed again in PBS for 5 minutes. This was followed by two washes in 2% FBS in 1×PBS, each for 5 minutes. Tissue sections were circled with a wax pen to retain liquid on top of the tissue by preventing leakage. Donkey blocking buffer (Table 2) was applied directly to the tissue within the wax barriers for 1 hour at room temperature. Primary antibody anti-GLUL (Rabbit) was diluted 1 in 800 in donkey blocking buffer and incubated on the tissue overnight at 4°C. The next day slides sequentially in PBST (0.1% Tween-20 in 1 x PBS) for 1, 5, 10 and 15 minutes. Secondary antibody donkey anti-rabbit Alexa-Fluor 594 were diluted 1:800 in PBS and incubated on tissue for 1 hour at room temperature, followed by sequential washes in PBST for 1, 5 and 10 minutes. A second blocking step was performed using rabbit blocking buffer (Table 2) for 1 hour at room temperature. Rockland GFP-conjugated antibody was diluted 1 in 800 in PBS and incubated on tissue at 4°C overnight. Slides were then washed 3 times in 0.1% (v/v) Tween-20 in 1×PBS for a duration of 1, 5 and 10 minutes sequentially. Slides were washed in PBS then incubated for 5 minutes in 0.8μg/mL DAPI in PBS for nuclear staining. Slides were again washed in PBS then dried carefully around tissue. Approximately 3 drops (~100 μL) of Gelvatol mounting medium were applied to each slide, followed by placement of a coverslip. Slides were left to dry in the dark for at least ten minutes prior to imaging or stored at 4°C in the dark until required. Imaging occurred shortly after staining to minimise loss of signal over time.

2.9.7 Microscopy and image analysis

Images acquired using Zeiss Axio Imager A1 microscopes. Zeiss ZEN 3.8 software was used for microscope setup, image capture and image processing. Image analysis was performed using QuPath (214). Scripts used for image analysis can be found in Appendices (Section 6.1.3).

2.10 Statistical analysis and figure generation

GraphPad Prism (10.4.1) software was used to undertake all statistical analyses and generate graphs. Figures generated using BioRender.com.

`

3. Genome editing murine *Pah* locus using CRISPR/Cas9 system

3.1 Introduction

Phenylketonuria (PKU) is an attractive candidate for gene therapy because it is a monogenic disease with a clear therapeutic target, reducing blood phenylalanine concentrations. Treatment efficacy can be monitored non-invasively via dried blood spot analysis. Despite dietary management, there remains a substantial unmet need, as PKU can cause irreversible neurotoxicity if uncontrolled in early development, and significantly impacts quality of life in adults.

In this chapter, we describe the development of a universal CRISPR/Cas9 based strategy to insert a codon optimised *PAH* cDNA into intron 1 of the endogenous locus, preserving native transcriptional regulation. To achieve this, we employed a HITI-based repair strategy which favours correct donor orientation through iterative excision and reintegration. In this design, a Cas9 guide targeting intron 1 generates a double-stranded break in the genome, and a donor cassette flanked by matching guide sequences is also excised from its plasmid backbone. Correctly oriented integration disrupts the guide sites and therefore preventing recutting, whereas the reverse-oriented integration reconstitutes the target sites, allowing re-excision and repeated opportunities for correct insertion.

Implementing this approach requires delivery of both editing machinery and donor template to hepatocytes. AAV vectors are widely used in gene therapy because they efficiently transduce cells and deliver DNA cargo, with particularly strong liver tropism, but have a limited packaging capacity. To overcome this, we adopted a dual-vector system: one delivering SaCas9 and its guide RNA (174) and the other carrying the donor cassette. We

`

employed the AAV8 capsid, which efficiently transduces hepatocytes in mice and non-human primates following intravenous administration (215,216).

To enable a universal editing strategy, we further incorporated bicistronic expression elements into the donor design. Bicistronic constructs allow two proteins to be expressed from a single mRNA transcript, most commonly achieved using either internal ribosome entry sites (IRES) or a 2A peptide sequence between two sequences. IRES elements allow for cap-independent translation initiation, binding the ribosome and translation initiation factors through complex secondary structures (217). However, they are relatively large (>500 bp) and the level of expression of the downstream gene is often as low as 10-20% of the upstream gene (218), albeit with the advantage that each protein is translated independently (219).
In contrast, 2A peptides are short sequences (~60 bp) that cause ribosomal skipping at the C-terminal glycine-proline peptide bond of the 2A, leading to the release of the upstream protein with a short C-terminal tag, and the initiation of translation of the downstream protein beginning with a proline residue at the N terminus (220). Although these residual 2A amino acids may influence protein function, 2A elements generally provide stronger downstream expression compared to IRES. However, they are not without limitations as ribosomal drop-off or incomplete skipping can occasionally occur, producing either only the upstream product or an unintended fusion protein (221). Among the various 2A sequences, the porcine teschovirus-1 2A (P2A) sequence shows the highest skipping frequency (222) and was therefore chosen for use in this study to maximise downstream PAH expression.

3.2 Results

3.2.1 PKU mouse model colony establishment and phenotype characterisation

We successfully established a PKU mouse model (Pah$^{\text{enu2}}$) colony from heterozygous breeders imported from The Jackson Laboratory. Following genotyping of F1 progeny, heterozygous *Pah*$^{+/\text{enu2}}$ females were paired with homozygous *Pah*$^{\text{enu2/enu2}}$ males for all subsequent breeding (Figure 3.1 A).

After establishing the colony, we characterised the baseline phenotype of the model. Homozygous *Pah*$^{\text{enu2/enu2}}$ mice were distinguishable from their heterozygous and wildtype littermates by their lighter brown coat colour (Figure 3.1 B). *Pah*$^{\text{enu2/enu2}}$ mice exhibited significantly elevated blood phenylalanine levels and exhibited significantly lower blood tyrosine levels compared to their wildtype littermates (Figure 3.1 C-D), consistent with previous reports for this model (223).

Dietary modulation demonstrated that blood phenylalanine could be further increased by providing high-protein chow (Figure 3.1 E) or supplementing phenylalanine in drinking water (Appendices Figure 6.2). However, Phe-supplemented water proved impractical due to its low solubility and variable intake. In addition, prolonged high-protein feeding has been associated with physiological stress and an increased incidence of neurological symptoms including seizures, in Pah$^{\text{enu2}}$/$^{\text{enu2}}$ mice. For these reasons, high protein chow was used only as a brief, 3-day challenge later in the study to transiently elevate phenylalanine levels before specific experimental endpoints. Unless this short dietary intervention is stated, all mice were maintained on standard chow throughout the study.

`

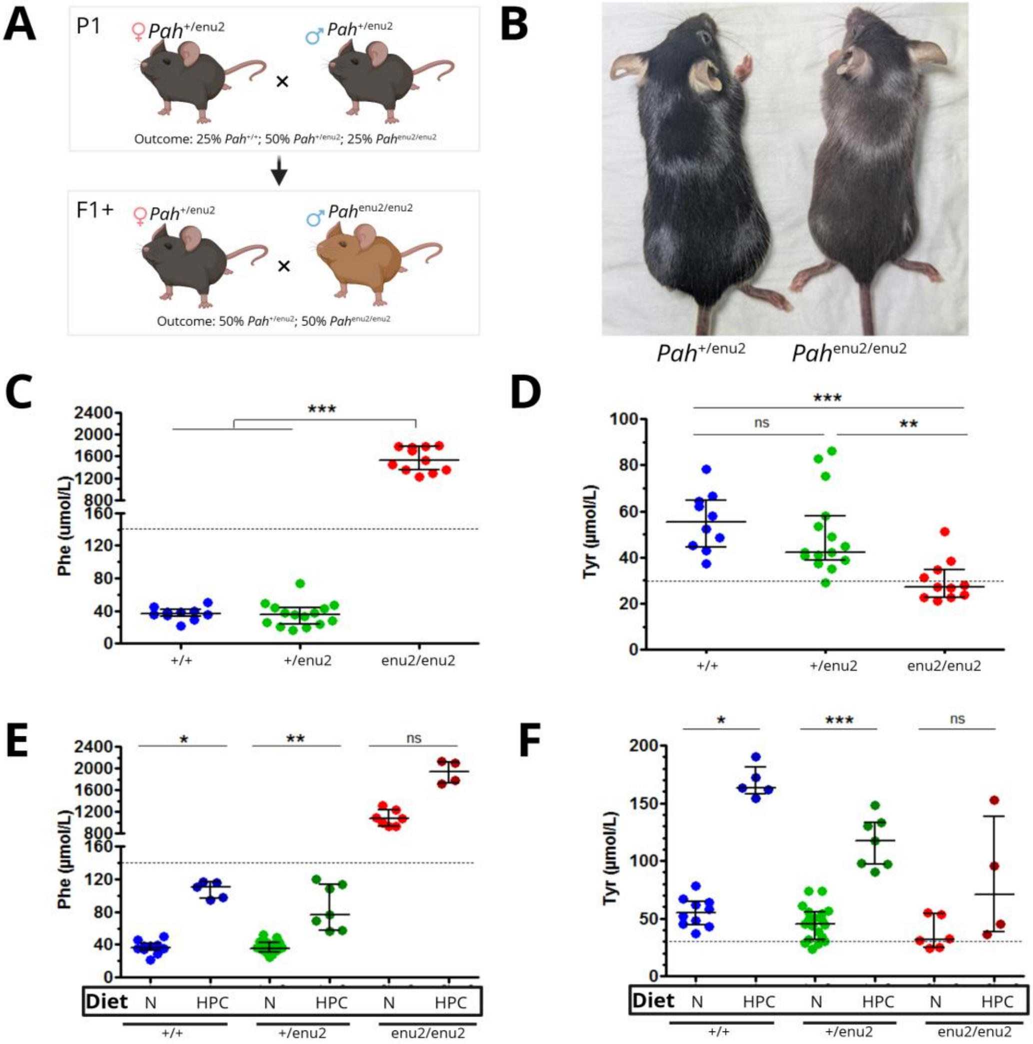


Figure 3.1: Establishment of Pah^{enu2} mouse model for PKU

(A) Breeding regime to establish the colony from JAX imported heterozygous Pah+/enu2 pairs (P1). Progeny were used to set up experimental breeding regime, heterozygous Pah+/enu2 females with homozygous Pahenu2/enu2 males (F1 onwards). (B) Representative image of coat colour phenotype of Pahenu2/enu2 mice. Mutant mice exhibit light brown coat in contrast to their black furred heterozygous littermates due to local tyrosine deficiency impeding production of melanin in the fur. Blood (C) phenylalanine and (D) tyrosine across genotypes on regular chow determined after 4 hours fast. Blood (E) phenylalanine and (F) tyrosine in heterozygotes compared to Pahenu2/enu2 mutants on regular chow (N) compared to high protein chow (HPC). Bars represent median and interquartile range. Statistical significance was assessed using the Kruskal-Wallis test followed by Dunn's multiple comparisons post hoc test. $p < 0.05$ considered significant. ns = not significant; * $p<0.05$; ** $p<0.01$; ***$p<0.0001$. Dashed lines represent the upper limit of healthy range of phenylalanine and the lower limit of the healthy range of tyrosine. JAX, The Jackson Laboratory; WT, wildtype; Phe, Phenylalanine; Tyr, Tyrosine.

3.2.2 Mutation-independent genome editing strategies to treat PKU

Two mutation-independent genome editing strategies were developed to treat PKU by targeted insertion of partial or full human *PAH* cDNA sequences into murine *Pah* intron 1 (Figure 3.2).

The endogenous *Pah* locus was targeted rather than a genomic safe harbour to preserve physiological regulation of PAH and to avoid the need for an exogenous promoter in the donor construct. *PAH* expression may vary with developmental or metabolic state, which would be lost if driven by a heterologous promoter. Furthermore, targeting a safe-harbour locus would leave mutant *Pah* alleles intact, allowing continued production of mutant polypeptides that could form mixed tetramers with the therapeutic protein and reduce functional enzyme activity. Overexpression of PAH could also lead to competition for the essential cofactor BH_4, which is shared with tyrosine hydroxylase and tryptophan hydroxylase, leading to the reduction of the activity of these other enzymes and potentially causing downstream metabolic consequences. Insertion into an intronic region of the endogenous locus places the therapeutic cassette under the control of the native *Pah* promoter, providing an important safety advantage, as off-target integration events are unlikely to result in ectopic expression in the absence of an active promoter.

Two constructs were evaluated. The first donor vector, referred to as the “Hybrid” approach, contains codon-optimised human *PAH* exons 2-13 downstream of a splice site and followed by a polyadenylation signal. This design enables splicing with native murine exon 1, forming a fusion protein to address mutations in exon 2-13 only. It is therefore intended for use in PKU patients with mutations downstream of exon 1, representing 95% of known pathogenic mutations.

The second construct, referred to as the “Universal” approach contains codon-optimised human PAH exons 1-13 preceded by a splice site and a P2A peptide sequence, followed by a polyadenylation signal. The P2A peptide sequence enables polycistronic expression from a single promoter by causing ribosomal skipping during translation, allowing for correction of all coding mutations in PAH. While P2A peptide sequences have been shown to facilitate efficient ribosomal skipping *in vivo*, occasional ribosomal drop-off or incomplete skipping may occur, potentially leading to unintended peptide products.

The first intron of *Pah* was selected as the target site for several reasons. Its size permits deep intronic editing, reducing the risk that CRISPR/Cas9-induced indels disrupt essential splice donor or acceptor sites. Second, targeting an early intron maximised the number of pathogenic mutations addressed by the Hybrid approach. For the Universal approach, targeting intron 1 minimises the size of any upstream truncated gene product which increases the likelihood that resultant mRNA undergoes nonsense-mediated decay rather than potentially interfering with enzyme function. Choosing a non-coding region also reduced the potential impact of unintended on-target events, such as large insertions, deletions, or incidental AAV integration, which could otherwise disrupt residual Pah function.

For these reasons, targeting intron 1 of the endogenous *Pah* locus was considered the most appropriate strategy for achieving mutation-agnostic correction while maintaining physiological regulation and minimising safety risks.

`

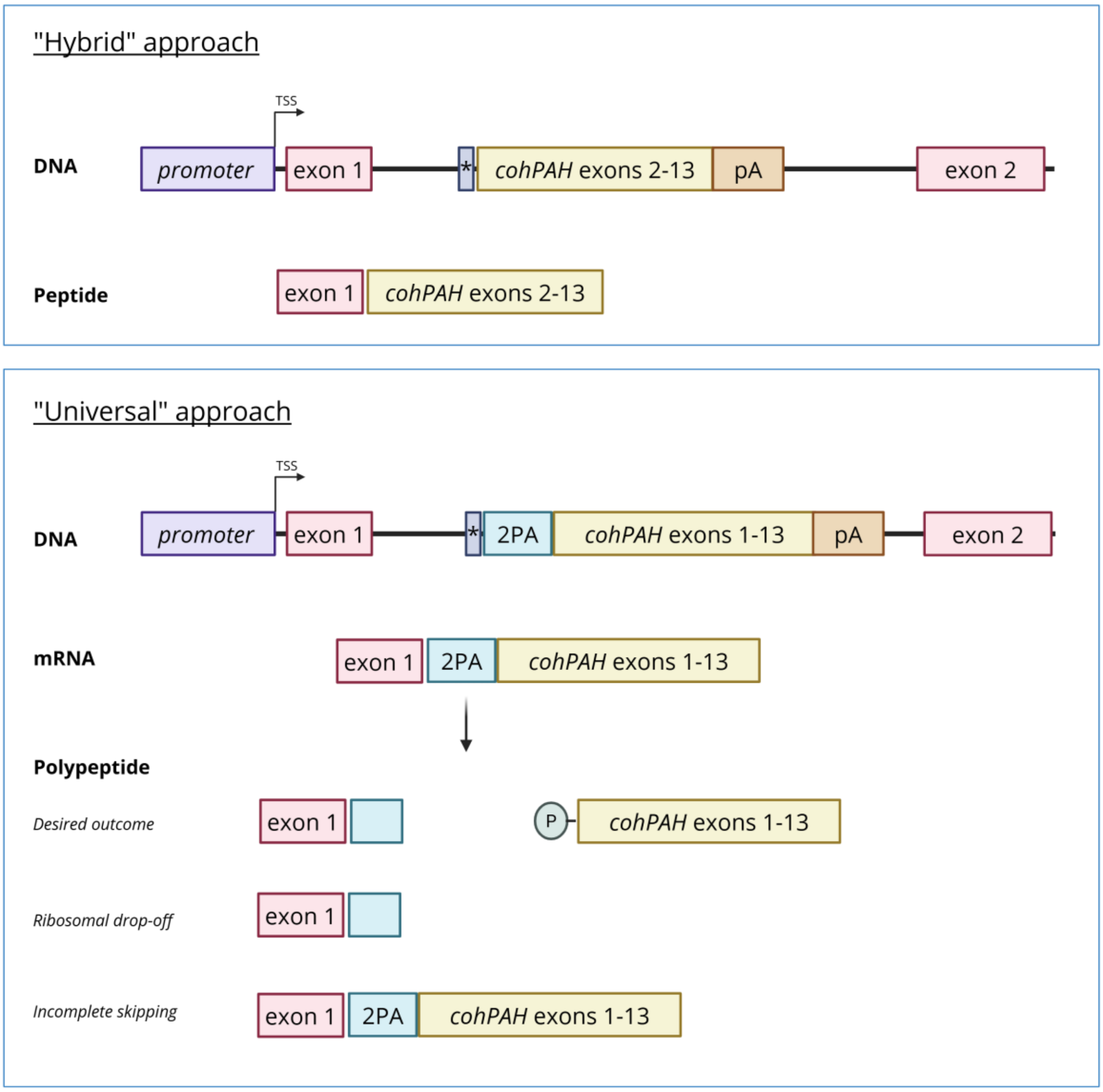


Figure 3.2: Diagram of gene-editing approaches at intron 1 of the murine *Pah* locus.

The “Hybrid” approach (top) contains a canonical splice branch and acceptor sequence (denoted by *), followed by *cohPAH* exons 2-13 cDNA and a BGH poly(A). This creates a fusion peptide containing murine exon 1 *Pah* with *cohPAH* exons 2-13, which goes on to form tetrameric PAH protein. The “Universal” approach (bottom) contains the same canonical splice branch and acceptor site (denoted by *), P2A peptide sequence, *cohPAH* exons 1-13 cDNA and a BGH poly (A). P2A peptide sequence efficiently induces ribosomal skipping, preventing the formation of a peptide bond at the C-terminus of the P2A. This results in the translation of the first exon of *mPah* carrying a short C-terminal P2A tag and *cohPAH* exon 1-12 with a proline at the N-terminus. Although ribosomal skipping with P2A peptides is usually efficient *in vivo*, occasional ribosomal drop-off or incomplete skipping may lead to unintended peptide products. TSS, Transcriptional Start Site; *cohPAH*, codon optimised human PAH; p(A), BGH polyadenylation signal sequence.

`

3.2.3 Design of SaCas9 sgRNAs targeting *Pah* intron 1

Staphylococcus aureus Cas9 (*Sa*Cas9) was selected for targeting intron 1 of the *Pah* gene primarily due to its compact size, which, in contrast to the larger *Streptococcus pyogenes* Cas9 (*Sp*Cas9), can be packaged into a single AAV vector (174). Furthermore, *Sa*Cas9 may offer increased editing precision and reduced risk of off-target edits, due to its more stringent protospacer adjacent motif (PAM) sequence requirements (173). A panel of guide RNA target sites (21 bp) with flanking *Sa*Cas9 PAM sequences (NNG(A/G)(A/G)T) were identified within *Pah* intron 1 using the CRISPOR *in silico* tool (224). CRISPOR ranks all identified gRNA and PAM sequences first by predicted specificity (MIT Specificity Score; (225)) and secondly by predicted cleavage efficiency (226).

To prevent potential disruption of normal splicing, all candidate sites were located at least 2 kb away from the endogenous splice acceptor and splice donor sites (Figure 3.3). Four candidate sgRNAs and their associated PAM sequences were selected based on specificity and efficacy scores, as well as minimal potential off-target regions and avoidance of common single nucleotide polymorphisms (SNPs) that may preclude certain populations from access (Table 3.1).

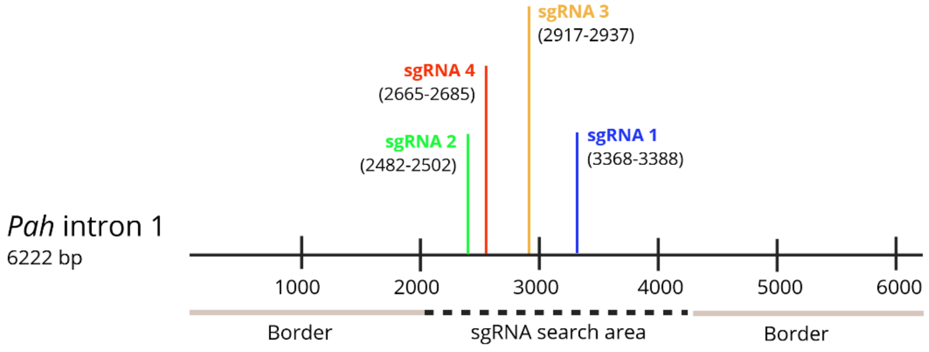


Figure 3.3: Location of selected sgRNA targets across murine Pah intron 1.

Guide RNA and PAM sequences were selected outside of 2 kb boundary from neighbouring exons to avoid potential disruption of splicing machinery, protecting the coding sequence of a potentially hypomorphic allele. *Pah*, Phenylalanine Hydroxylase; PAM, protospacer adjacent motif; sgRNA, single guide RNA.

`

Table 3 1: Candidate SaCas9 sgRNAs and their respective PAM sequences†.

#	sgRNA sequence 5'-3'	Strand	PAM	Specificity	Efficiency	Off-target for 0-1-2-3-4 mismatches
1	TAGCACATGCCGATATGCACC	+	AAGAAT	100	73	0-0-0-0-0
2	CCGGGATTGGCTAGGTAAACC	+	TTGAAT	99	69	0-0-0-0-3
3	GGGCCATGTTACATCATAAAT	+	GTGAAT	93	69	0-0-0-0-8
4	CAGGCTTTCTTACGTGTTAGC	+	TGGAAT	92	52	0-0-0-1-7

† Selected *in silico* using the web-based CRISPOR tool (224).

3.2.4 Evaluation of sgRNAs for SaCas9 in murine cell lines *in vitro*.

Candidate guides were assessed *in vitro* using two different murine immortalised cell lines, NIH3T3 and BW35. NIH3T3 cells are immortalised mouse fibroblast cells while BW35 cells are murine T-cell lymphoma cells. Neither cell line express *Pah* endogenously. Individual candidate guides were cloned into pX601 plasmid (174), which contains SaCas9 under a CMV promoter and sgRNA1, 2, 3 or 4 under a ubiquitous U6 promoter. SaCas9-sgRNA constructs. Cell lines were transfected using both lipofectamine and calcium phosphate methods.

The efficiency of the selected candidate sgRNAs were evaluated by PCR across the predicted CRISPR/SaCas9 cleavage site using extracted genomic DNA from cells. PCR amplicons were Sanger sequenced, and the resulting chromatograms were analysed using the online Inference of CRISPR Edits (ICE) tool (227). ICE performs sequence chromatogram decomposition by aligning sequence data from edited samples to a reference (untreated control), allowing the estimation of the proportion of alleles containing insertions and deletions (indels) at the target site. A higher percentage of indels at the cleavage site indicates a greater proportion of error-prone repair via the NHEJ pathway, which repairs double-stranded DNA breaks (228). As such, indel percentage is frequently used as a measure of CRISPR/Cas9 cleavage efficiency (229). Anticipated indels across the cut site were not

`

detected using either approach, despite confirming that the transfection was efficient using a plasmid containing a CMV-GFP expression cassette as a control. As the model had then arrived and had excess stock while establishing the colony, we moved straight into *in vivo* guide screening.

3.2.5 Evaluation of sgRNAs for SaCas9 in murine model of PKU *in vivo*.

Candidate guides were assessed *in vivo* in the murine PKU mouse model, *Pah*$^{\mathrm{enu2}}$. Individual candidate guides were cloned into pX602 plasmid (211), which contains SaCas9 under a liver specific TBG promoter and sgRNA1, 2, 3 or 4 under a ubiquitous U6 promoter (Figure 3.4 A). CRISPR/SaCas9-sgRNA constructs were delivered as rAAV2/8 vector via tail vein injection at a dose of 5×10^{11} vg/mouse to *Pah*$^{+/\mathrm{enu2}}$ mice at 5 weeks (Figure 3.4 B). Mice were sacrificed 4 weeks later and livers extracted for DNA analysis. Heterozygous mice were used to minimise animal waste and reduce project timelines while establishing the model colony. Candidate sgRNA targeted regions within intron 1, identical regions in both genotypes and as such should not affect guide performance.

Again, indel frequency was assessed using PCR. The percentage of indels showed a clear candidate to move forward with, with sgRNA1 being the only guide to produce indels, with an average indel rate of ~16% (Figure 3.4 C). To determine if the percentage of indels identified in genomic DNA isolated from the liver of treated animals were due to variation in vector transduction, the number of Cas9 vector copies per diploid genome (VCN per dg) was quantified using ddPCR (Methods 2.9.1). Differences in guide performance were not attributed to variation in liver copy number, suggesting that indel rates were due to the sgRNA sequence itself (Figure 3.4 D).

`

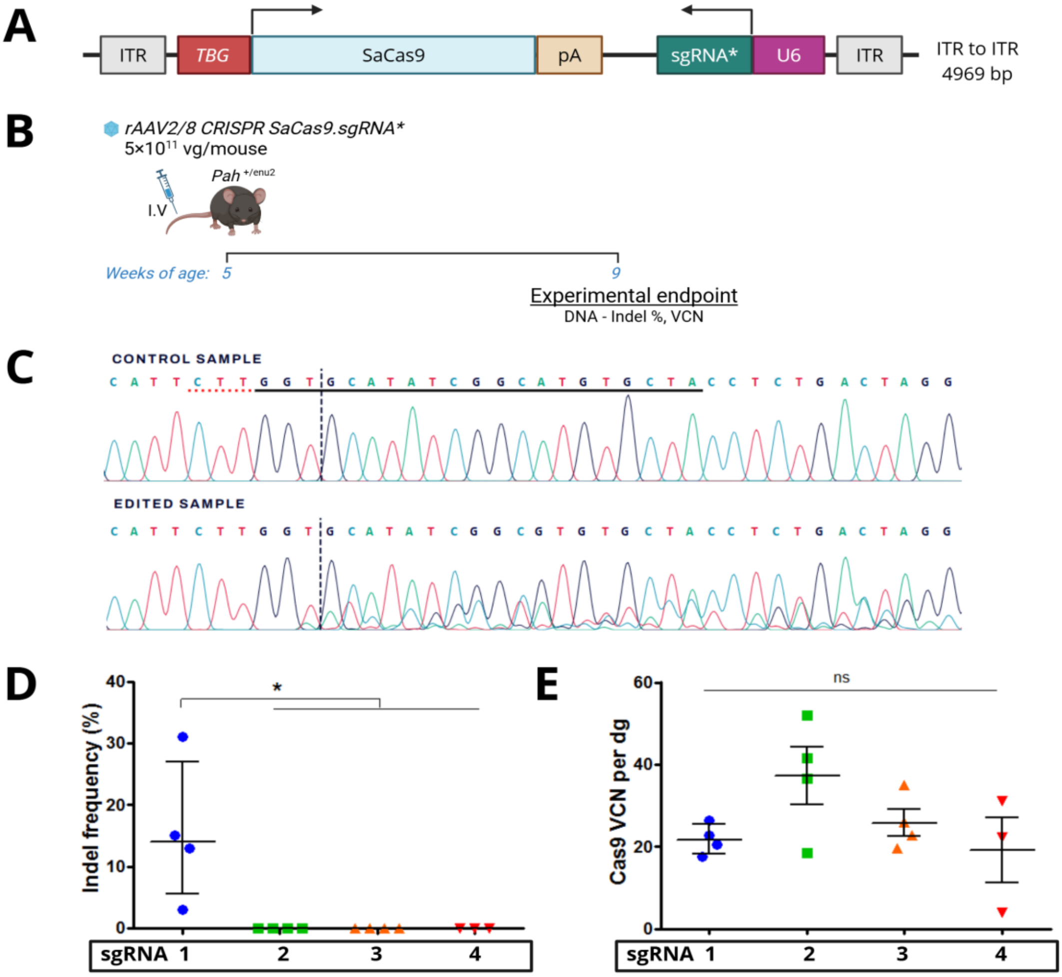


Figure 3.4: Evaluation of candidate SaCas9 sgRNAs *in vivo*.

(A) Configuration of the pX602 plasmid containing SaCas9 and sgRNA used. ”*” represents each of the four candidate guides tested. (B) Experimental overview. Heterozygous *Pah*$^{+/enu2}$ mice received intravenous delivery of rAAV2/8 SaCas9-sgRNA vector at 5×10^{11} vg/mouse at five weeks of age (four mice per group). Mice were sacrificed four weeks following vector delivery and livers were extracted for DNA analyses. Sample size (n>4) used per treatment group. (C) Synthego ICE *in silico* tool sequence chromatogram decomposition of comparing data from edited sample aligned to a control sample. Horizontal black underlined region represents the guide sequence. Horizontal red underline is the PAM site. Vertical black dotted line represents the actual cut site. (D) Analysis of indel % in each treated mouse by candidate sgRNA. Amplicons generated by PCR across the cleavage site for each sgRNA were Sanger sequenced. Reads were aligned to reference sequence and proportion containing insertions and deletions. (E) Vector copy number per diploid genome. Analysis was performed using ddPCR comparing *Sa*Cas9 to Albumin in treated mice liver genomic DNA. Bars represent median and interquartile range. Statistical significance was assessed using the Kruskal-Wallis test followed by Dunn’s multiple comparisons post hoc test. $p < 0.05$ considered significant. ns = not significant; * p<0.05. Dg; diploid genome; Indel, insertions and deletions; ITR, inverted terminal repeat; pA, BGH polyadenylation signal; *Sa*Cas9, *Staphylococcus aureus* Cas9 nuclease; VCN, vector copy number.

`

3.2.6 Design of a dual rAAV vector system to enable targeted insertion of donor-DNA vector when paired with CRISPR-Cas9 vector

A dual rAAV vector system was developed to facilitated targeted insertion of donor DNA sequences using the HITI strategy. Two donor vectors encoding either GFP or human-codon optimised PAH (hcoPAH) cDNA were designed following the hybrid and universal approaches described in Section 3.3.4 (Figure 3.5 A-D). Each donor sequence was flanked on each side by the sgRNA1 recognition sites and PAM sequence, enabling CRISPR/Cas9 mediated cleavage. This configuration allows preferential integration in the correct orientation, as reveres integration reconstitutes the sgRNA sites and can be subject to Cas9 re-excision (Figure 3.5 B).

Reporter (GFP) and therapeutic (PAH) vectors were evaluated in parallel. GFP provided a convenient readout, since protein expression could be readily quantified using immunohistochemistry and flow cytometry. In contrast, PAH protein cannot be distinguished in this model, since available antibodies detect both wild-type and mutant protein alike, and as such GFP reporter vectors were used to infer performance of PAH therapeutic vectors. The GFP or PAH coding sequences were codon-optimised for human expression *in silico* (GenScript, Nanjing, China) to improve translation efficiency by selecting codon usage based on the specific selected (230). Restriction enzyme recognition sites for PacI and SpeI were excluded to facilitate easy cloning into the self-complementary AAV backbone (pAAV2-LPhOTC BB2 supplied by A/Prof Samantha Ginn). Constructs included either the full-length hPAH cDNA, P2A-hPAH exons 2-13, EGFP cDNA or P2A-EGFP. The human PAH sequence was selected for the therapeutic donors to maximise clinical translatability, as prior studies have demonstrated that the human sequence can cross complement murine *Pah* (141).

Synthesised donor sequences (GenScript) were cloned into the scAAV backbone, verified by Sanger sequencing and packaged into rAAV2/8 capsids.

For *in vivo* evaluation, heterozygous $Pah^{+/\mathrm{enu2}}$ mice were treated with GFP donors, while homozygous $Pah^{enu2/\mathrm{enu2}}$ mice were used for hcoPAH therapeutic donors. Heterozygous $Pah^{+/\mathrm{enu2}}$ mice were injected with 1×10^{12} rAAV2/8 CRISPR/SaCas9-sgRNA1 vector and 5×10^{11} rAAV2/8 GFP donor at 8 weeks and harvested 4 weeks later for analysis (Figure 3.5 E). Homozygous $Pah^{enu2/\mathrm{enu2}}$ mice were injected with 1×10^{12} rAAV2/8 CRISPR/SaCas9-sgRNA1 vector alongside 5×10^{11} rAAV2/8 hcoPAH donor at 8 weeks and harvested approximately 8.5 weeks later for analysis (Figure 3.5 F), consistent with reports that phenylalanine correction plateaus by ~6 weeks in AAV editing studies (192). Control groups received only the donor vectors (5×10^{11} rAAV2/8). Blood was collected at baseline and longitudinally post injection for phenylalanine and tyrosine analysis. Vector dose and ratio was determined by previous research in the lab (213). Heterozygotes were used to test GFP constructs to minimise animal wastage and reduce project timelines.

`

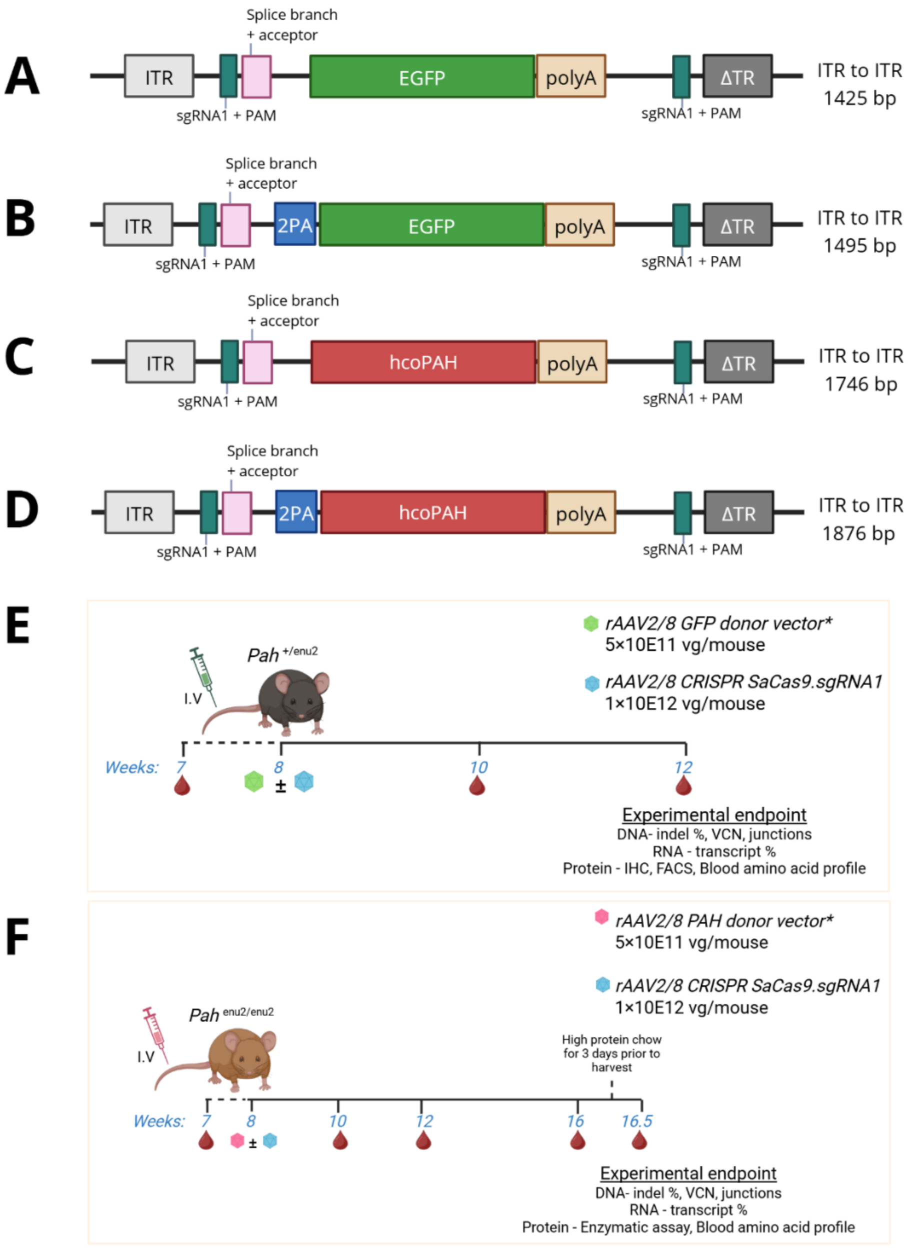


Figure 3.5: Design and testing of mutation-independent genome editing approach using dual AAV CRISPR/SaCas9 and donor repair templates *in vivo*.

Configuration of the donor vector plasmids used in this project. (A) Hybrid GFP scAAV, (B) Universal GFP scAAV, (C) Hybrid cohPAH scAAV and (D) Universal cohPAH scAAV. ΔTR (dark grey) allows for the generation of scAAV vectors. (E) Experimental overview for GFP cohort. *Pah*$^{+/enu2}$ mice received intravenous delivery of rAAV2/8 GFP donor vector at 5×10^{11} vg/mouse either alone or alongside rAAV2/8 SaCas9-sgRNA vector at 1×10^{12} vg/mouse at 8 weeks of age. Mice were sacrificed 4 weeks following vector delivery and livers were extracted for DNA, RNA and protein analyses. Blood was taken at 7,10 and 12 weeks of age. Sample size (n≥4) used for each treatment group. (F) Experimental overview for hcoPAH cohort. *Pah*$^{enu2/enu2}$ mice received intravenous delivery of rAAV2/8 PAH donor vector at 5×10^{11} vg/mouse either alone or alongside rAAV2/8 SaCas9-sgRNA vector at 1×10^{12} vg/mouse at 8 weeks of age. Mice were sacrificed at 16.5 weeks following vector delivery, with a 3-day protein challenge prior to harvest. Livers were extracted for DNA, RNA and protein analyses. Blood was taken at 7,10,12,16 and 16.5 weeks of age. Sample size (n≥4) used for each treatment group. EGFP, enhanced green fluorescent protein; hcoPAH, human codon optimised PAH; Indel, insertions and deletions; ITR, inverted terminal repeat; ΔTR, inverted terminal repeat containing terminal resolution site deletion; 2PA, 2A peptide 2PA; pA, BGH polyadenylation signal; scAAV, self-complementary AAV; *SaCas9*, *Staphylococcus aureus* Cas9 nuclease; VCN, vector copy number.

3.2.7 Dual rAAV CRISPR/Cas9-sgRNA1 and GFP donor leads to GFP expression in a high proportion of hepatocytes *in vivo*

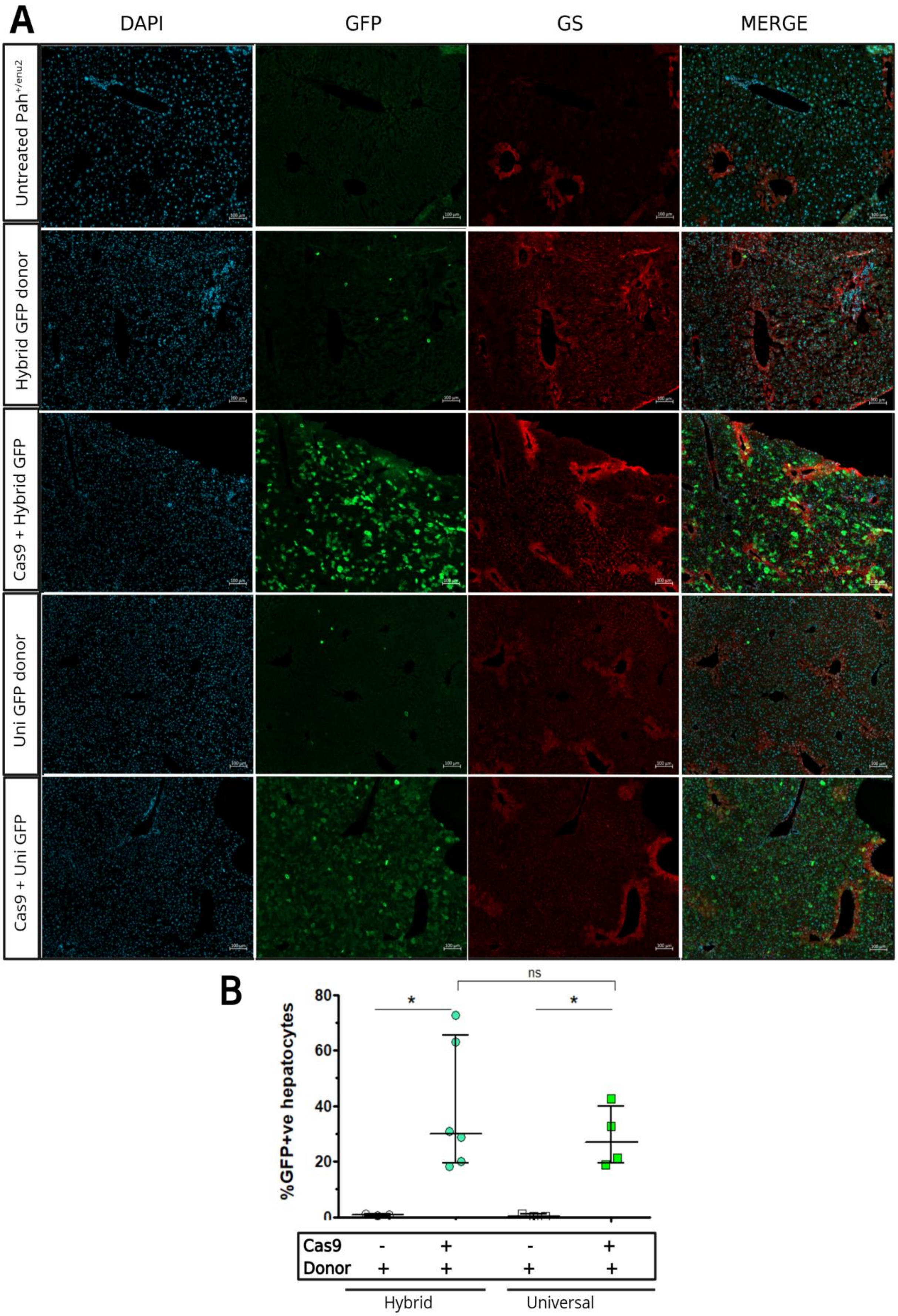
Expression of the GFP transgene was assessed using immunofluorescence and flow cytometry analysis. Cryopreserved liver tissue from mice that received either dual CRISPR/SaCas9-sgRNA1 and GFP donor vectors or the GFP donor vectors alone were co-stained with GFP and glutamine synthetase (GS). GS staining was used as a zonation marker, labelling 1-2 hepatocytes surrounding the central vein (231).

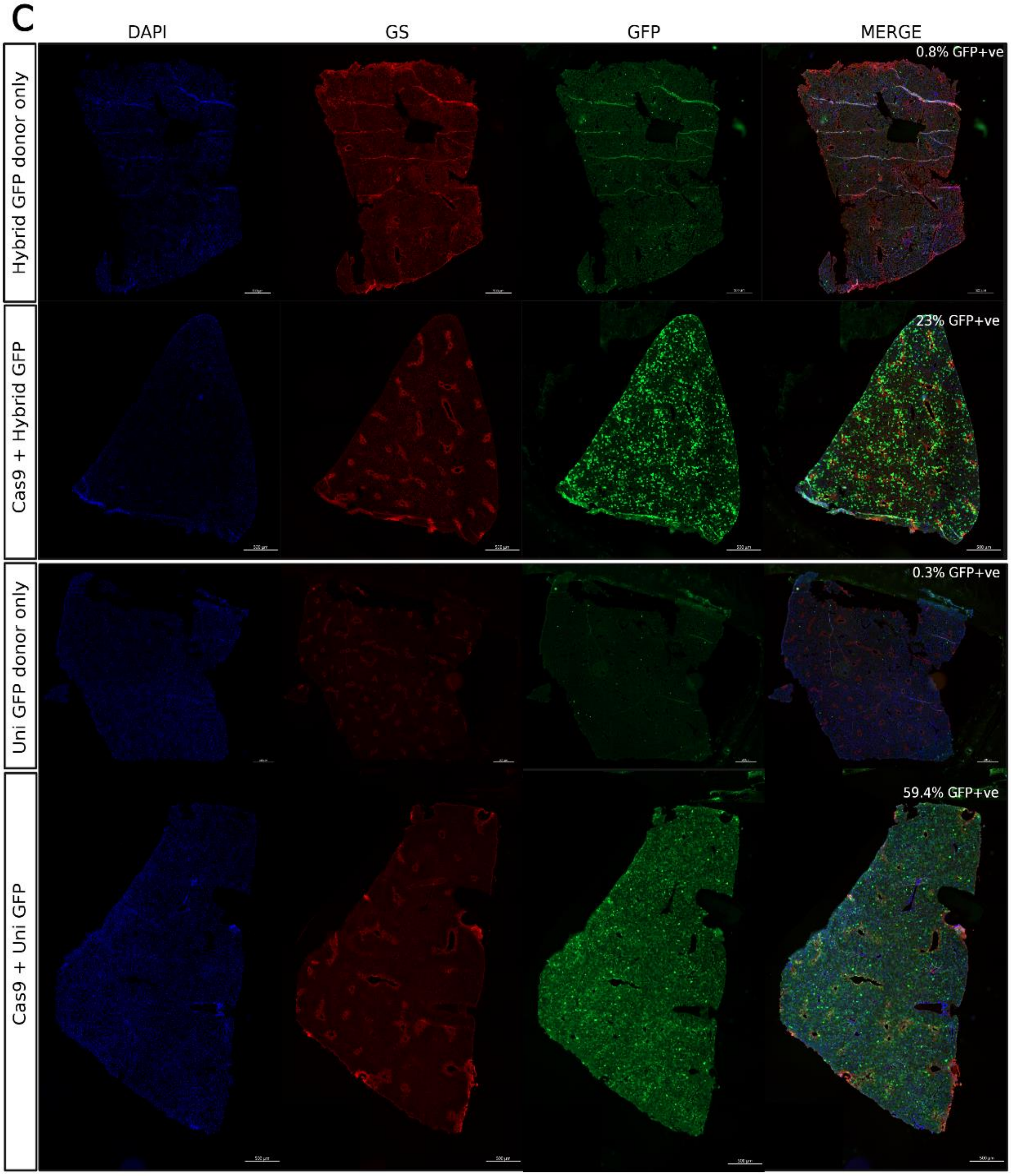
Mice treated with both CRISPR/Cas9-sgRNA1 and GFP donor vectors exhibited high but variable levels of GFP expression across the liver, while mice receiving the donor vector alone displayed minimal to no detectable GFP signal (average 0.51%, range: 0-0.98%) (Figure 3.6 A). Mice that received the CRISPR/Cas9-sgRNA1 alongside the hybrid GFP donor vector showed an average of 39% GFP-positive hepatocytes (range: 18.27%-72.66%, SD: 23.11%) whereas those receiving the universal GFP donor vector averaged 43% GFP-positive hepatocytes (range: 18.81%-42.55%, SD: 10.95%). No significant difference in the proportion of GFP-positive hepatocytes was observed between the two dual-treated groups (Figure 3.6 B). However, GFP intensity was brighter in the dual hybrid GFP group compared to the dual universal GFP group. Co-staining with GS indicated a consistent enrichment of GFP-positive hepatocytes in the pericentral region across liver sections (Figure 3.6 C).

To complement immunofluorescence data, hepatocytes were isolated via liver perfusion for flow cytometric quantification of GFP expression. Due to suboptimal perfusion and poor cell recovery, only a subset of samples could be analysed. Nonetheless, the available data was consistent with the results obtained using immunofluorescence. High levels of GFP expression were observed in mice receiving both vectors, while those receiving the donor alone saw negligible levels (Figure 3.6 D and E). Among dual treated mice, those receiving

`

the hybrid GFP donor vector exhibited a mean GFP fluorescence intensity that was two-folder higher than those receiving the universal donor vector (Figure 3.6 F). These differences were expected because the universal donor construct employed a P2A peptide sequence to co-express GFP, which can result in reduced downstream gene expression due to incomplete ribosome skipping or ribosome drop-off (221). Finally, there was a strong correlation in the mice assessed by both immunohistochemistry and flow cytometry which added confidence to our analyses ($r^2 = 0.9942$) (Figure 3.6 G).


A
DAPI
GFP
GS
MERGE
Untreated $Pah^{+/enu2}$
Hybrid GFP donor
Cas9 + Hybrid GFP
Uni GFP donor
Cas9 + Uni GFP
100 µm
B
%GFP+ve hepatocytes
0
20
40
60
80
*
*
ns
Cas9 - + - +
Donor + + + +
Hybrid
Universal


C
DAPI
GS
GFP
MERGE
Hybrid GFP donor only
0.8% GFP+ve
Cas9 + Hybrid GFP
23% GFP+ve
Uni GFP donor only
0.3% GFP+ve
Cas9 + Uni GFP
59.4% GFP+ve
500 µm

`

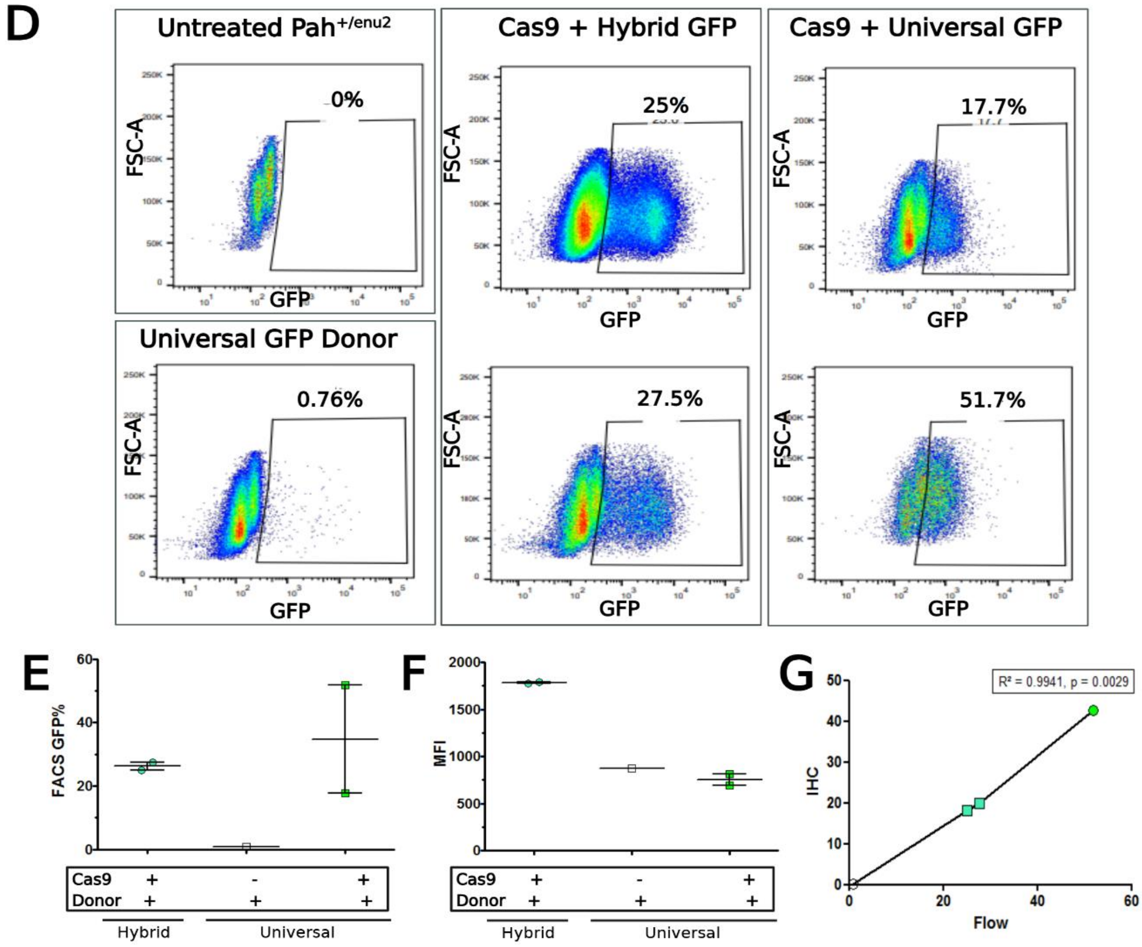


Figure 3.6: Analysis of GFP expression in dual and donor-only treated mice reveals GFP expression in a high proportion of hepatocytes *in vivo*.

(A) Immunofluorescence images of liver sections from *Pah*$^{+/enu2}$ mice treated with rAAV2/8 CRISPR/SaCas9-sgRNA1 and/or rAAV2/8 GFP donor vectors and untreated controls. Liver sections were stained with antibodies against GFP (green) and glutamine synthetase (GS) (red). Nuclei were stained with DAPI. Scale bar= 100 µm. (B) GFP-positive cell counts performed by custom QuPath macro (Supplementary X). Bars represent median and interquartile range. Statistical significance was assessed using one-tailed Mann Whitney U test. * $p<0.05$. (C) Mosaic immunofluorescence images of entire liver sections from Pah+/enu2 mice treated with rAAV2/8 CRISPR/SaCas9-sgRNA1 and/or rAAV2/8 GFP donor vectors and untreated controls. Liver sections were stained with antibodies against GFP (green) and glutamine synthetase (GS) (red). Nuclei were counter stained with DAPI. Scale bar= 500 µm. (D) Flow cytometry analysis of proportion of GFP positive cells (GFP%) by treatment group. (E) Mean fluorescent intensity (MFI) per treatment group. FACS and data analysis performed by Cindy Zhu using FlowJo. The gating strategy used forward side scatter (FSC-A) versus GFP fluorescence (GFP) to determine the percentage of GFP-positive hepatocytes, as described in Supplementary X. (F) A strong correlation was observed between the percentage of hepatocytes expressing GFP determined using IHC and GFP expression measured by flow cytometry.

`

3.2.8 Quantification of perfect HITI genomic junctions showed modest levels of targeted integration into *Pah* intron 1

Genomic junctions between the endogenous murine *Pah* locus and inserted donor vectors were quantified using dPCR. Junction-specific primers were designed to detect HITI-mediated integration in either the correction or the reverse orientation (Figure 3.7 A). A PCR spanning the CRISPR cut site was also performed to assess locus disruption. Successful targeted insertions yield large amplicon that exceeds the efficient amplification range of the assay and are therefore not directly detected by dPCR, but rather indirectly as a dropout in recoverable PCR products. All the PCR amplicons were validated by Sanger sequencing (Figure 3.7 B-C).

Both correct and reverse genome-donor junctions were detected exclusively in dual-treated mice (Figure 3.7 D and F). This confirms that CRISPR/Cas9-mediated cutting facilitated donor sequence integration at the predicted location. In dual GFP-treated mice, junction levels were comparable between hybrid and universal donor groups (Figure 3.7 D). Correct junctions averaged 5.02% (range: 2.1-9.23%) in the hybrid GFP group and 6.59% (range: 3.03-10.18%) in the universal GFP group, while reverse junctions were detected at 4.28% (range: 2.09-8.15%) and 4.44% (range: 2.27-8.12%) respectively. A significant reduction in recoverable PCR products spanning the cut site was also observed, consistent with targeted insertions disrupting amplification across the locus (average entire GFP cohort 78.98%; range: 64.86-91.89%) (Figure 3.7 E).

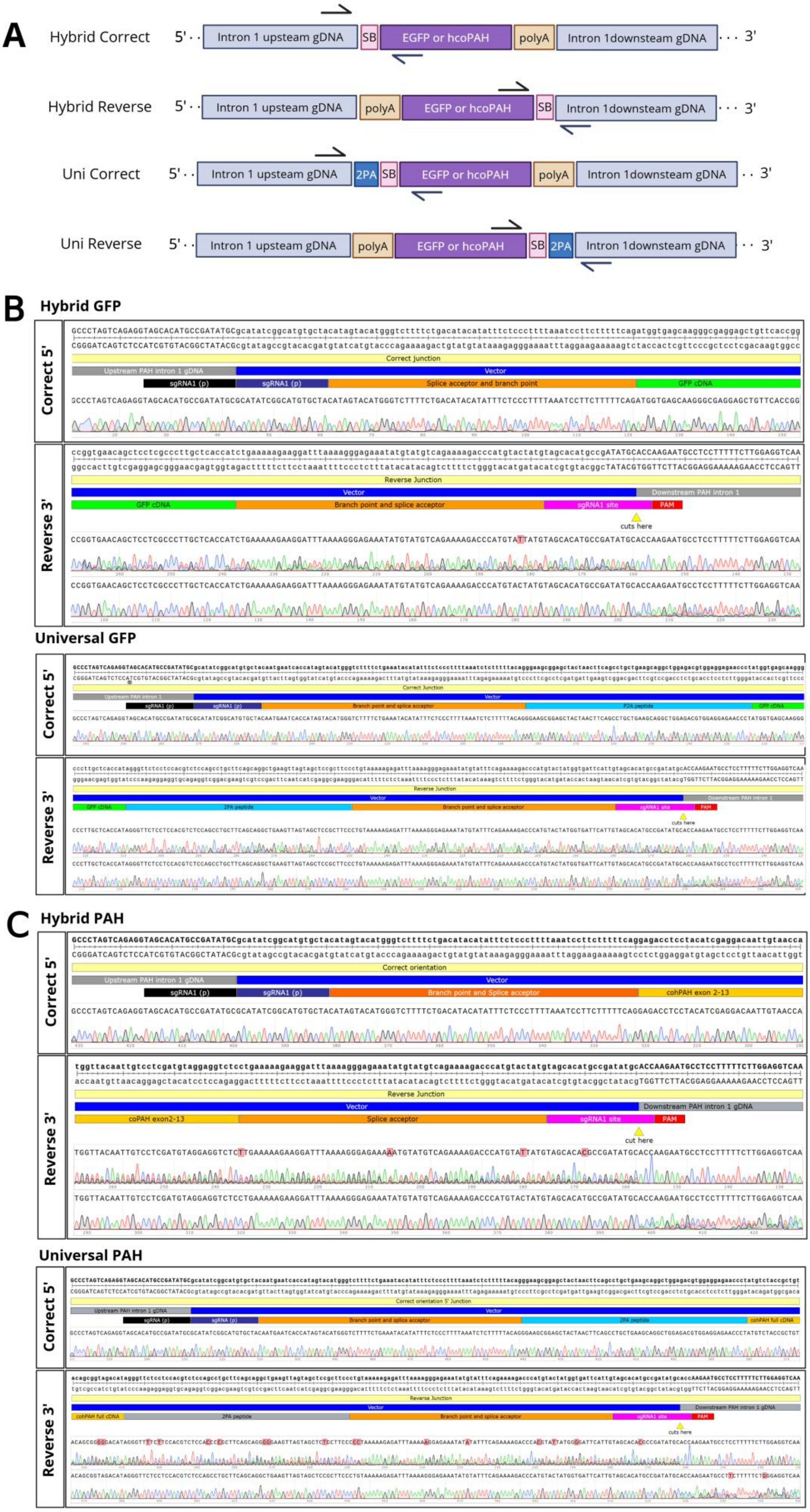
Interestingly, the combined proportion of detected junctions and unedited loci amplicons did not approach 100%, suggesting that that were events at the target site that could not be amplified using the PCR strategy. These events may include (i) large deletions that disrupt or

`

remove primer binding sites or (ii) integration of full concatemer or partial AAV vector sequences forming imperfect HITI junctions that are not captured by the current primer design (Ginn et al., 2025). Notably, the number of genome- GFP junctions detected by dPCR was markedly lower than the proportion of GFP-positive hepatocytes observed by immunohistochemistry and flow cytometry.

Dual PAH-treated mice exhibited comparable levels of genome-donor junctions (Figure 3.7 F). The hybrid hcoPAH group showed correct junctions at a mean of 5.96% (range: 2.95-7.9%) and reverse junctions at 2.53% (range: 1.63-4.16%). The universal hcoPAH group has correct junctions at 4.16% (range: 1.23-6.15%) and reverse junctions at 2.65% (range: 1.36-3.48%). Unlike the GFP groups, there was no significant reduction in the recoverable PCR products spanning the cut site, however there is a trend downwards with a large range (67.77-100%) (Figure 3.7 G). Notably, mouse #213 which was treated with the universal PAH donor, saw only ~68% recovery across the cut site, which was better than the average of the entire GFP cohort mice.

Across all dual treated groups, the proportion of correct junctions was significantly higher than reverse junctions (Figure 3.7 H, $p<0.01$ for both), indicating a modest but consistent bias towards integration in the correct orientation. In GFP mice harvested 4-week post treatment, correction:reverse ratio was approximately 1.3:1, while in PAH mice 8.5 weeks after treatment this was 1.6:1. This bias did not change significantly over time, although the interval before cohorts was only 4.5 weeks and overall junction recovery was low.


A
Hybrid Correct
5'
Intron 1 upsteam gDNA
SB
EGFP or hcoPAH
polyA
Intron 1downsteam gDNA
3'
Hybrid Reverse
Uni Correct
2PA
Uni Reverse
B
Hybrid GFP
Correct 5'
Correct junction
Upstream PAH intron 1 gDNA
Vector
sgRNA1 (p)
Splice acceptor and branch point
GFP cDNA
Reverse 3'
Reverse Junction
Downstream PAH intron 1
Branch point and splice acceptor
sgRNA1 site
PAM
cuts here
Universal GFP
P2A peptide
2PA peptide
C
Hybrid PAH
Correct orientation
Branch point and Splice acceptor
cohPAH exon 2-13
coPAH exon2-13
Splice acceptor
Downstream PAH intron 1 gDNA
cut here
Universal PAH
Correct orientation 5' Junction
sgRNA (p)
cohPAH full cDNA

`

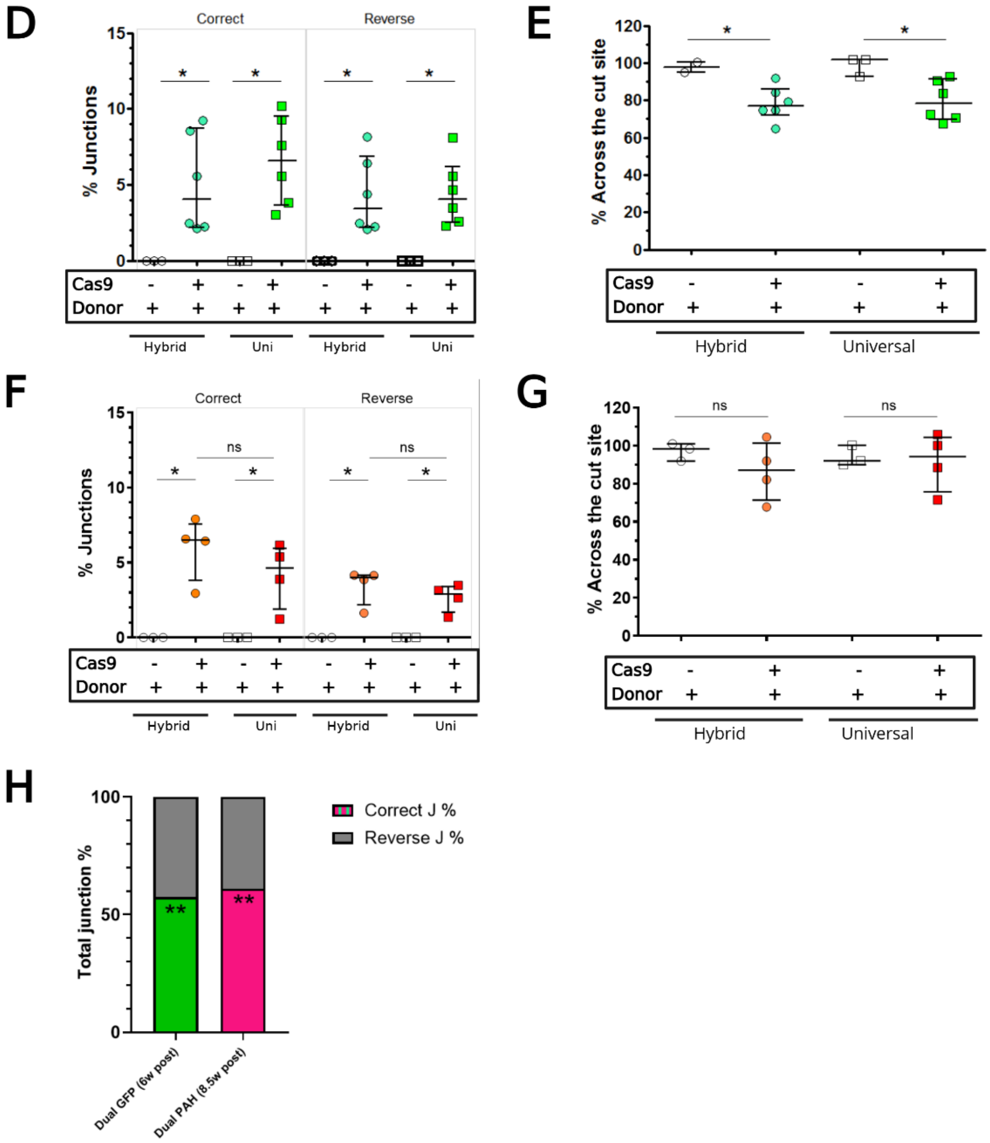


Figure 3.7: Detection of HITI genomic junctions in dual-treated mice by dPCR.

(A) Location of primers used to hybrid genomic-donor junctions in both the correct and reverse orientations. (B-C) Representative PCR traces confirming the specificity of amplicons used for quantification. (D) Proportion of genomic-GFP donor junctions detected in the correct and reverse orientation by dPCR, normalised to albumin. GFP donor treated mice were harvested 4 weeks after treatment. (E) Proportion of PCR products recoverable across the sgRNA cut site in GFP donor treated cohort. (F) Proportion of genome-hcoPAH donor junctions detected in the correct and reverse orientation by dPCR, normalised to albumin. cohPAH donor treated mice were harvested 8.5 weeks after treatment. (G) Proportion of PCR products recoverable across the sgRNA cut site in PAH donor treated cohort. Bars represent median and interquartile range in (D-G). (H) Mean percentage of correct and reverse junctions relative to total junctions isolated in each cohort. Stacked bars represent group mean. Statistical significance was assessed using either one tailed Mann-Whitney U test (D-G) or paired t-test (H). *p* < 0.05 considered significant. ns = not significant; * p<0.05; **p<0.01

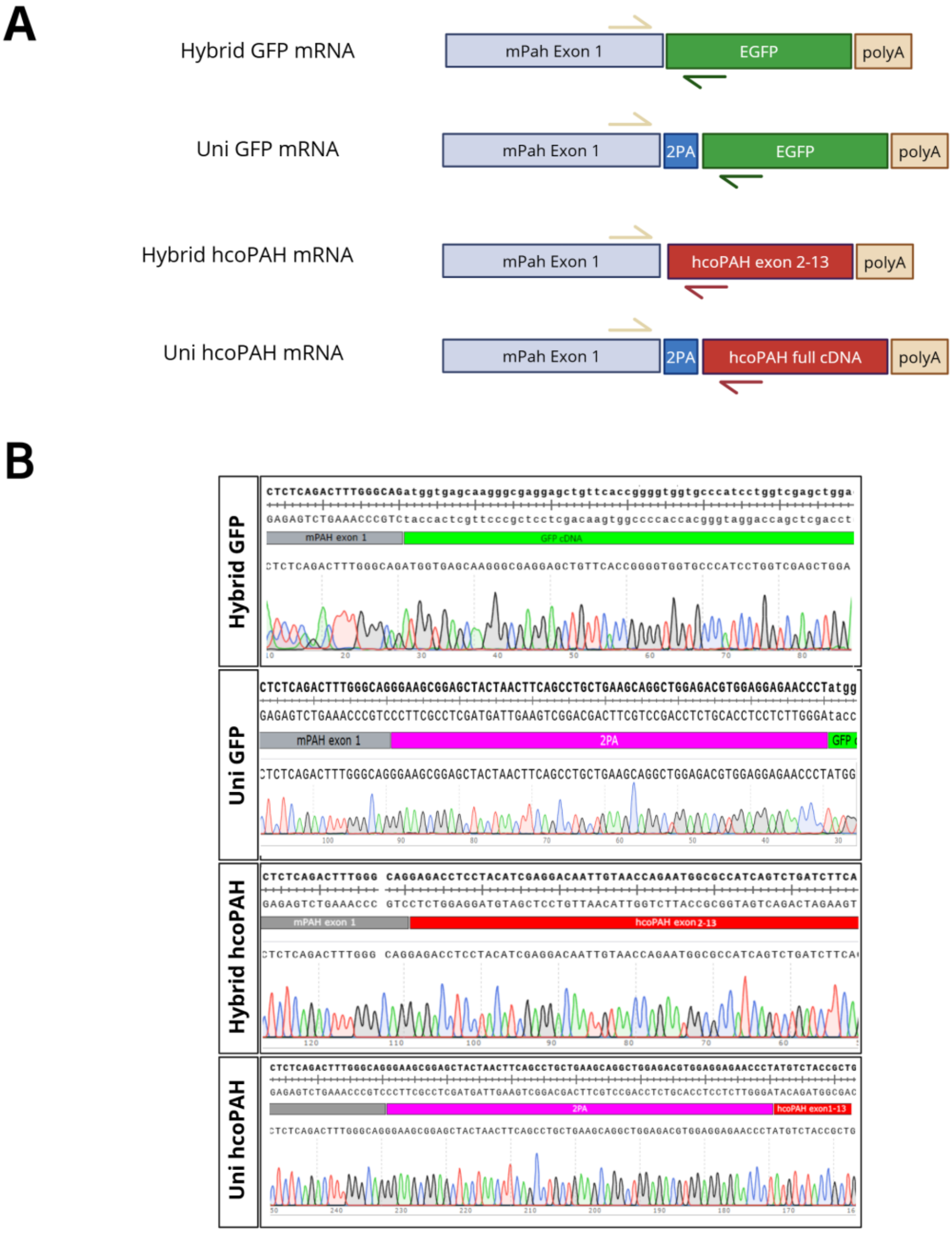
3.2.9 Analysis of transcript in dual treated mice confirms successful on-target editing

Another measure of targeted integration performance was the quantification of hybrid transcript levels relative to endogenous *Pah* mRNA levels. RNA was extracted from the liver of vector treated mice and reverse transcribed to generate cDNA libraries. Quantification of hybrid transcript was assessed by dPCR using a universal forward primer targeting murine *Pah* exon 1 paired with reverse primers specific to the 5' end of GFP, hcoPAH or the endogenous murine *Pah* exon 3 (Figure 3.8A) and normalised to the amount of *hprt1* cDNA in each sample. Amplification specificity was confirmed by sequencing (Figure 3.8 B). Both Hybrid GFP and Hybrid PAH transcript were detected exclusively in dual-treated mice, consisted with CRISPR mediated on-target integration. As both hybrid and endogenous transcripts are driven by the same promoter, an increase in hybrid transcript levels should coincide with a reduction in endogenous murine *Pah* transcript. However, no reduction in murine Pah mRNA levels was observed either across treatment groups or between untreated genotypes (Figure 3.8C).

Hybrid transcript abundance was calculated as a proportion of total transcript output from the locus (hybrid + endogenous *Pah*) and expressed as % hybrid transcript. Mice treated with dual GFP donor vectors exhibited comparatively high levels of GFP transcript. Dual hybrid GFP treated mice expressed an average of 27.5% GFP transcript from the locus (range:15.5-39%) while dual universal GFP treated mice saw an average of 29.5% (range: 16.6-44.5%). By contrast, PAH-donor treated mice saw only modest levels of hybrid transcript expression. Dual hybrid PAH treated mice saw an average of 6.2% PAH transcript from the locus (range:1.6-8.7%) whereas dual universal PAH treated mice saw a lower average of 3.8% (range: 1.2-6%). No significant difference between vector configurations (hybrid vs universal

`

approach) was observed. To state plainly, there was a significant difference in the hybrid transcripts detected between GFP and PAH donor vectors. GFP donor cohort ranged from 15.5-44.51%, while PAH donor cohort ranged from 1.6-8.8%.


A
Hybrid GFP mRNA
mPah Exon 1
EGFP
polyA
Uni GFP mRNA
mPah Exon 1
2PA
EGFP
polyA
Hybrid hcoPAH mRNA
mPah Exon 1
hcoPAH exon 2-13
polyA
Uni hcoPAH mRNA
mPah Exon 1
2PA
hcoPAH full cDNA
polyA
B
Hybrid GFP
CTCTCAGACTTTGGGCAGatggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctgga
GAGAGTCTGAAACCCGTCtaccactcgttcccgctcctcgacaagtggccccaccacgggtaggaccagctcgacct
mPAH exon 1
GFP cDNA
TCTCAGACTTTGGGCAGATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGA
10
20
30
40
50
60
70
80
Uni GFP
CTCTCAGACTTTGGGCAGGGAAGCGGAGCTACTAACTTCAGCCTGCTGAAGCAGGCTGGAGACGTGGAGGAGAACCCTatgg
GAGAGTCTGAAACCCGTCCCTTCGCCTCGATGATTGAAGTCGGACGACTTCGTCCGACCTCTGCACCTCCTCTTGGGAtacc
mPAH exon 1
2PA
GFP
TCTCAGACTTTGGGCAGGGAAGCGGAGCTACTAACTTCAGCCTGCTGAAGCAGGCTGGAGACGTGGAGGAGAACCCTATGG
100
90
80
70
60
50
40
30
Hybrid hcoPAH
CTCTCAGACTTTGGG CAGGAGACCTCCTACATCGAGGACAATTGTAACCAGAATGGCGCCATCAGTCTGATCTTCA
GAGAGTCTGAAACCC GTCCTCTGGAGGATGTAGCTCCTGTTAACATTGGTCTTACCGCGGTAGTCAGACTAGAAGT
mPAH exon 1
hcoPAH exon 2-13
TCTCAGACTTTGGG CAGGAGACCTCCTACATCGAGGACAATTGTAACCAGAATGGCGCCATCAGTCTGATCTTCA
120
110
100
90
80
70
60
Uni hcoPAH
CTCTCAGACTTTGGGCAGGGAAGCGGAGCTACTAACTTCAGCCTGCTGAAGCAGGCTGGAGACGTGGAGGAGAACCCTATGTCTACCGCTG
GAGAGTCTGAAACCCGTCCCTTCGCCTCGATGATTGAAGTCGGACGACTTCGTCCGACCTCTGCACCTCCTCTTGGGATACAGATGGCGAC
2PA
hcoPAH exon1-13
TCTCAGACTTTGGGCAGGGAAGCGGAGCTACTAACTTCAGCCTGCTGAAGCAGGCTGGAGACGTGGAGGAGAACCCTATGTCTACCGCTG
240
230
220
210
200
190
180
170

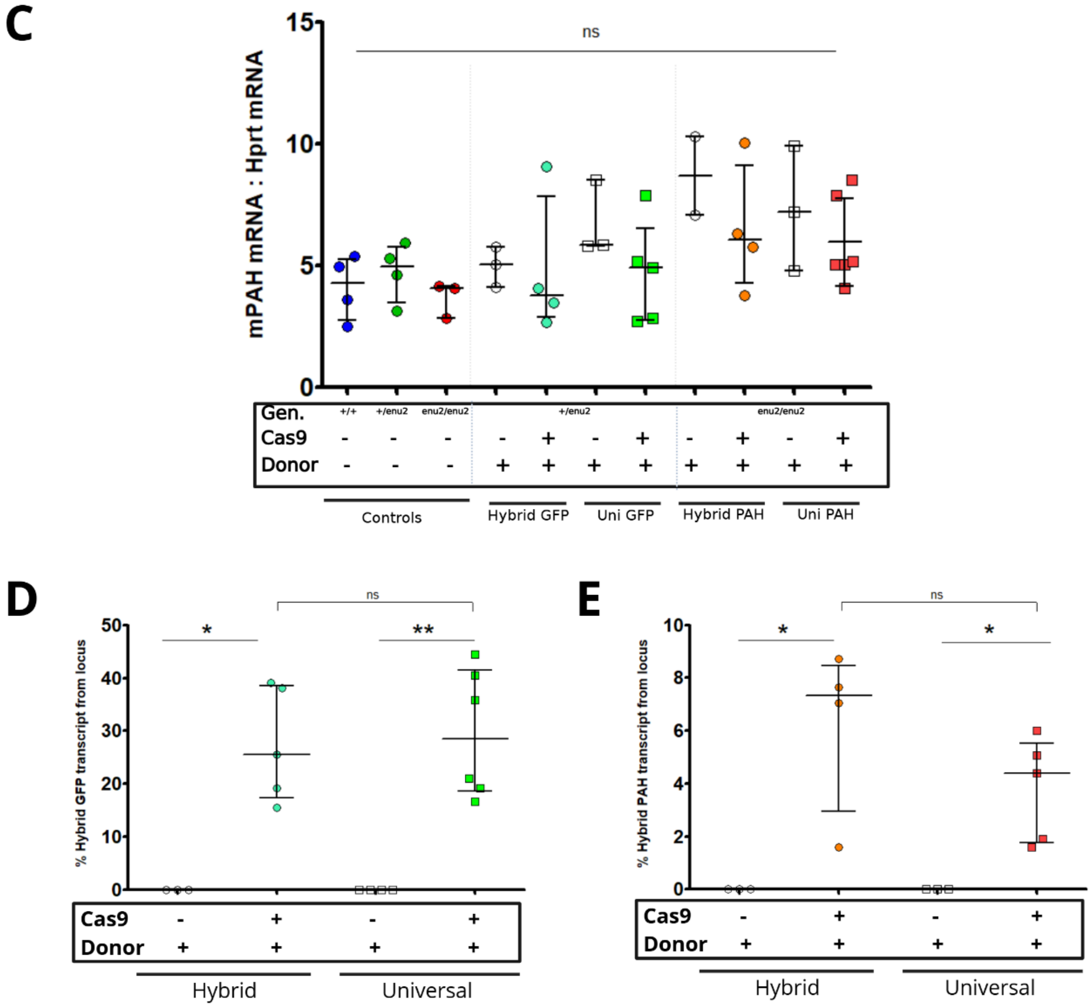


Figure 3.8: Quantification of hybrid Pah transcripts in mice following AAV delivery.

(A) Location of primers used to amplify the expected hybrid transcripts in the liver following treatment with both the Hybrid and Universal donor (GFP/PAH) approaches. (B) Sanger sequencing of PCR amplicons confirms the amplification of the hybrid transcripts. (C) Endogenous murine *Pah* mRNA expression in untreated control and treated mice relative to *Hprt1*. The proportion of (D) hybrid GFP transcripts and (E) hybrid *PAH* transcript relative to the total number of transcripts (total of endogenous and hybrid from *Pah* locus.

`

3.2.10 Indel and vector copy number analyses show no difference between GFP and PAH donor vector.

To understand why the level of hybrid transcripts was higher in animals receiving the donor vector containing the EGFP transgene in comparison to those receiving the hcoPAH transgene, a comprehensive set of molecular analyses were performed.

First, we examined indel frequencies across cohorts. As expected, indels were detected in only CRISPR treated mice, however considerable range was observed (0-18%), with many dual treated mice having no indels whatsoever. We observed no significant difference across the groups treated with different vector configurations or donors (Figure 3.9 A).

Next, we assessed vector copy number (VCN) to determine whether local vector dose could explain differences the differences in transgene expression. Cas9 vector copy number was only detected in dual treated cohorts, with no significant difference dual treated GFP and PAH groups (Figure 3.9 B). Similarly, analysis of donor vector levels revealed no significant differences across cohort (Figure 3.9 C-D).

Linear regression analyses were performed to examine the relationships between molecular measures and expression outcomes (Figure 3.9 E-F). In GFP-treated mice, the proportion of GFP-positive hepatocytes detected by IHC correlated significantly, though weakly with both Cas9 (r^2=0.7589, p=0.0048) and GFP VCN (r^2=0.6718, p=0.0128). Likewise, GFP transcript levels correlated with both Cas9 VCN (r^2=0.7165, p=0.0020). In PAH-treated cohorts, the proportion of PAH-hybrid transcript was significantly correlated with both PAH VCN (r^2=0.62, p=0.02) and Cas9 VCN (r^2=0.53, p=0.04) (Appendices Table 6.2).

`

These findings indicate that copy number of both Cas9 and donor vectors contribute to key steps in the targeted integration process-specifically, the efficient generation of double-stranded breaks and the availability of donor templates for integration at the target site. However, the correlations explained only a moderate proportion of the variance, suggesting that additional factors influence expression outcomes. It is also possible that the relationship between vector dose and functional output is non-linear, potentially requiring a threshold for robust activity.

`

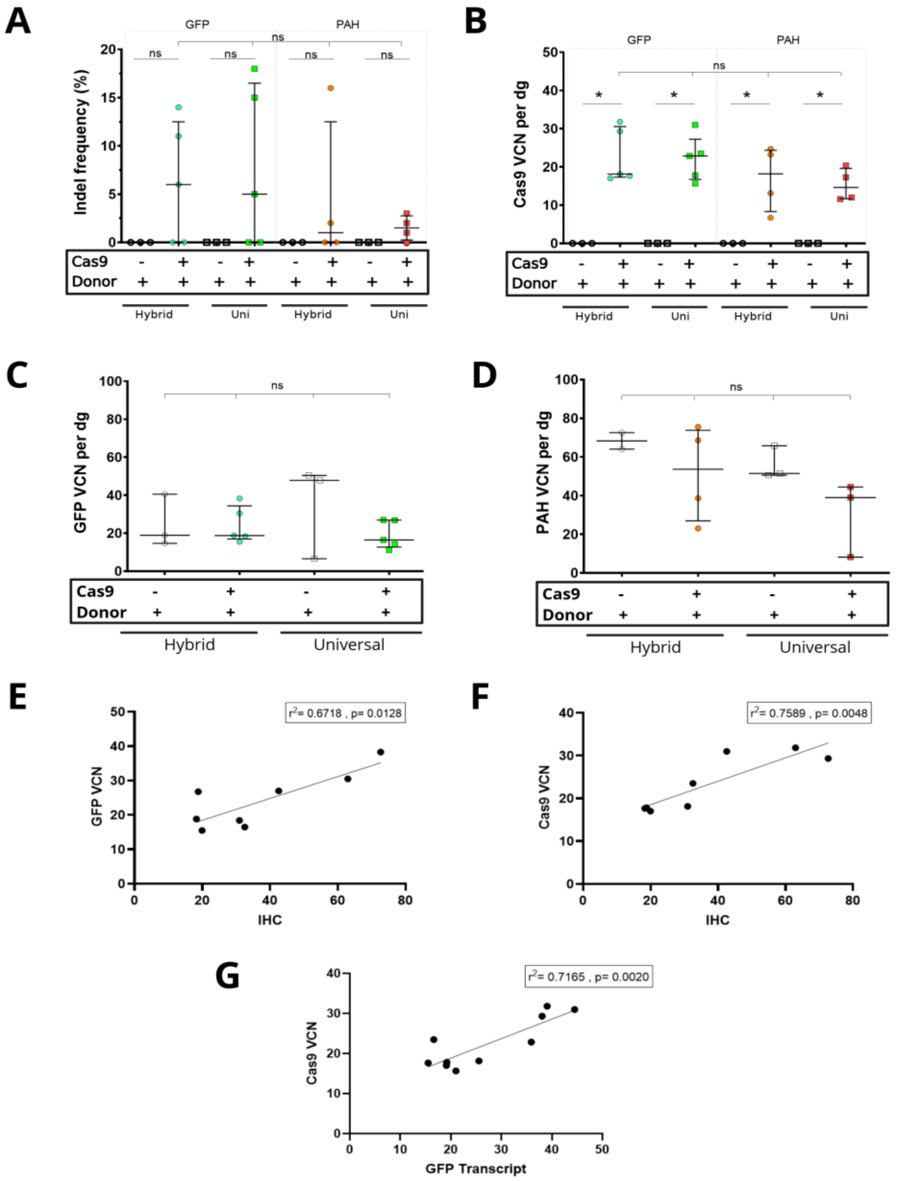


Figure 3.9: Indel and vector copy number analysis do not indicate difference in approach performance in both hybrid and universal approaches.

(A) Analysis of Indel % in each treated mouse either treated with donor vectors with or without CRISPR/Cas9-sgRNA1. Amplicons generated by PCR across the cleavage site for each sgRNA were Sanger sequenced. Reads were aligned to reference sequence (control mouse) and proportion containing insertions and deletions were quantified using the Synthego ICE in silico tool. (B) CRISPR/Cas9 vector copy number per diploid genome(C) GFP vector copy number per diploid genome. (D) PAH vector copy number per diploid genome. All VCN analysis was performed using ddPCR comparing target to Albumin in treated mice liver genomic DNA. Bars represent median and interquartile range. Statistical significance was assessed Statistical significance was assessed using one tailed Mann-Whitney U test. Linear regression analysis comparing (E) GFP IHC % vs GFP VCN (F) GFP IHC % vs Cas9 VCN (G) GFP transcript % vs Cas9 VCN. $p < 0.05$ considered significant. ns = not significant; * p<0.05. Dg; diploid genome; Indel, insertions and deletions; GFP; green fluorescent protein PAH; phenylalanine hydroxylase; *Sa*Cas9, *Staphylococcus aureus* Cas9 nuclease; VCN, vector copy number.

3.2.11 Blood phenylalanine levels were not significantly reduced despite reversion of light-fur phenotype in dual vector treated *Pah*$^{enu2/enu2}$ mice.

To assess functional efficacy of our intervention, blood phenylalanine and tyrosine levels were monitored throughout the experiments. Dried blood spots were collected via submandibular bleeds onto filter paper prior to treatment and at defined time points (7, 8, 10, 12 and 16 weeks) in PAH-donor treated cohorts (Figure 3.5 D). A terminal submandibular bleed was also performed following a three-day high protein dietary challenge prior to sacrifice. Quantification of blood phenylalanine and tyrosine levels were performed using LC-MS/MS by the Newborn Screening Laboratory at The Children’s Hospital at Westmead. Overall, no statistically significant change in blood phenylalanine or tyrosine was observed over time in any treatment group pre or post treatment and pre or post high protein challenge (Figure 3.10 A-B). However, some individual mice treated with dual CRISPR/Cas9-sgRNA and hcoPAH donor vectors did exhibit a modest reduction in blood phenylalanine levels. Specifically, mice #213 (dual hybrid GFP donor treated) and #181 (dual universal GFP donor treated) showed a 32.5 % and 49.4% decrease in blood phenylalanine respectively, on normal chow at 16 weeks compared to pre-injection levels (Figure 3.10 C). Only mouse #213 exhibited a blunted response to dietary protein load, maintaining stable phenylalanine levels accompanied by an increase in blood tyrosine levels (Figure 3.10 D). This may suggest that the additional dietary phenylalanine was successfully metabolised into tyrosine. Notably, mouse #213 had (i) the highest copy number for both the Cas9 and PAH donor vectors, (ii) the highest percentage of hybrid transcripts (8.9%), (iii) the highest percentage of indels (16%) and (iv) the second highest percentage of correct junctions detected (6.6%) of the entire PAH treated cohort. Mouse #181 had the highest number of hybrid transcripts (5.1%) and highest number of correct junctions (6.1%) for the dual universal donor cohort only.

`

Interestingly, a reversion of the light-fur phenotype was observed exclusively in dual treated mice (Figure 3.10 E). Untreated mice and those receiving donor only treatment retained the typical light brown fur, whereas dual treated mice gradually replaced the light fur with a dark brown fur as the hair moulted.

While these findings are encouraging, even the best performing mice exhibited blood phenylalanine levels that remain far from above clinical thresholds to confer a meaningful benefit for PKU patients. Further studies with optimised experimental design and larger sample sizes will be required to validate these preliminary findings, particularly given the baseline variability observed in blood phenylalanine levels.

`

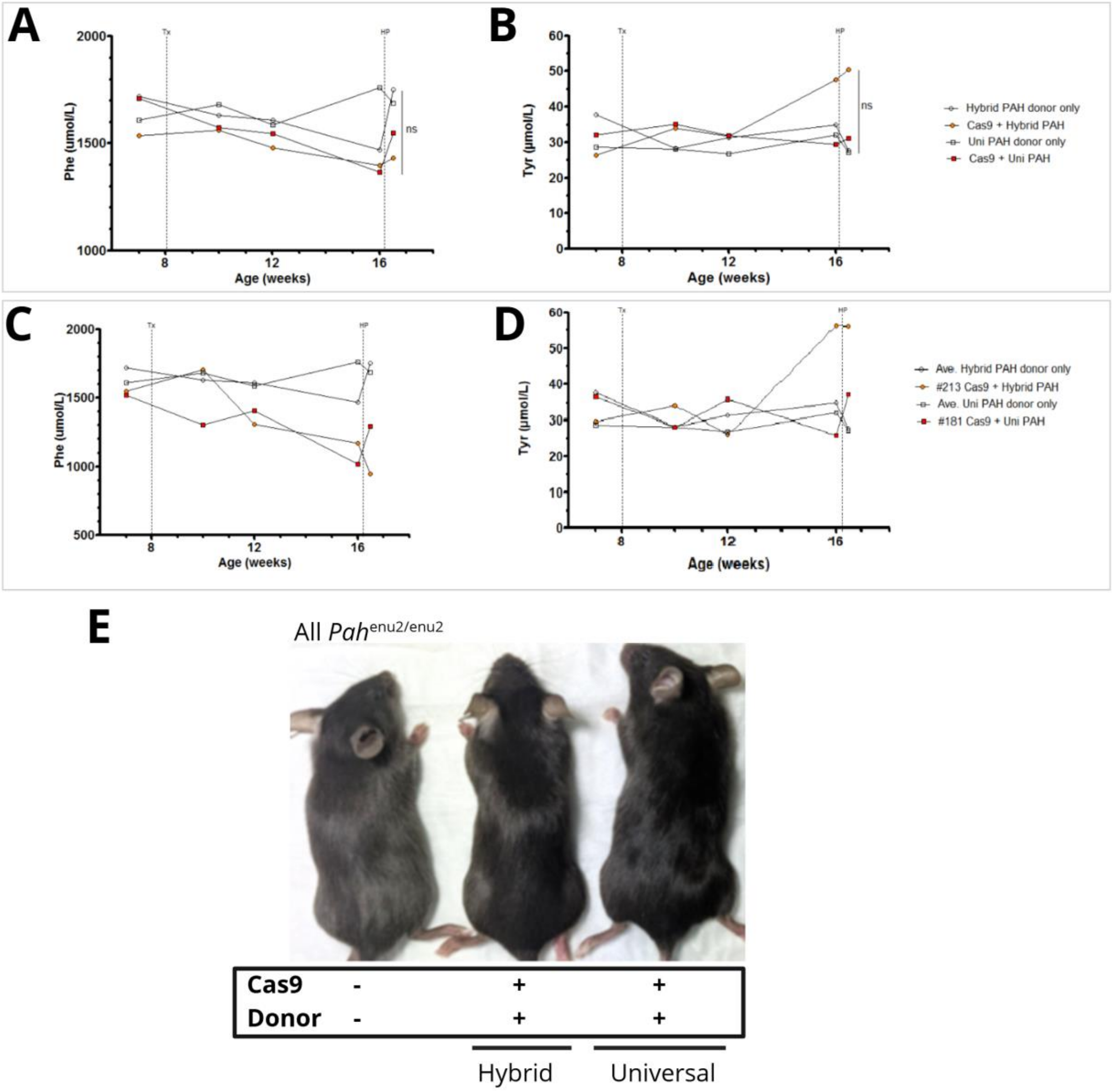


Figure 3.10: Blood phenylalanine levels were not significantly reduced despite reversion of light-fur phenotype.

Blood phenylalanine levels (A) and blood tyrosine levels (B) over time in cohort of mice treated with treated with rAAV2/8 CRISPR/SaCas9-sgRNA1 and/or rAAV2/8 hcoPAH donor vectors. Time of injection denoted by ‘Tx’. Mice were challenged with a high protein diet for three days prior to sacrifice, denoted by ‘HP’. Ns = non-significant ($p>0.05$). Blood phenylalanine levels (C) and blood tyrosine levels (D) over time in individual mice to illustrate variable response and downwards trend in dual treated mice. Time of injection denoted by ‘Tx’. Mice were challenged with a high protein diet for two days prior to sacrifice, denoted by ‘HP’. (E) Light brown fur colour typical of *Pah*$^{enu2/enu2}$ mutant mice is restored to black colour in dual treated mice only, signifying increases in local tyrosine enabling the production of melanin. More images available in Appendices Figure 6.3.

`

3.2.12 Development of an *in vitro* PAH enzymatic assay by LC-MS/MS

Concurrently, we sought to establish an *in vitro* PAH enzymatic assay as another measure of therapeutic performance. Total protein was extracted from snap-frozen, homogenised liver samples. 25 µg of protein was incubated for 15 minutes with stable isotope-labelled phenylalanine-d5 in the presence of the required cofactors tetrahydrobiopterin (BH4) and iron (Figure 3.11 A). The use of labelled phenylalanine and tyrosine allowed direct measurement of conversion while avoiding confounds from endogenous amino acids present in the system. The reaction was quenched with 10% formic acid containing 5µM tyrosine-d4, and the relative abundance of phenylalanine-d5 and tyrosine-d4 was quantified using LC-MS/MS. A standard curve was generated using known concentrations of phenylalanine and tyrosine as described by Heintz et al. (2012) (232). Sample preparation, LC-MS/MS and data analysis were all performed by Dr Mark Graham (CMRI, Biomedical Proteomics).

While both phenylalanine-d5 and tyrosine-d4 could be detected (Figure 3.11 B-C), the *in vitro* PAH activity assay requires further optimisation. Extracted ion chromatograms demonstrated clear peaks corresponding to the expected precursor ions for phenylalanine-d5 (m/z 171.1176) and tyrosine-d4 (m/z 186.1063), indicating that the system can detect both substrate and product in a protein extract background. The corresponding MS/MS spectra confirmed analyte identity through the expected precursor-to-fragment transitions, validating the detection method.

Liver lysates from wildtype mice were assayed both in the presence of BH4 (positive control) and absence of BH4 (negative control). No difference in tyrosine-d4 production was observed between these controls, and the measured levels of tyrosine-d4 were substantially lower than expected based on protein input (150 µM/L) (232). This suggest that while the LC-MS/MS

`

platform was successfully established, the enzymatic assay conditions were not yet optimised to support robust detection of PAH activity.

`

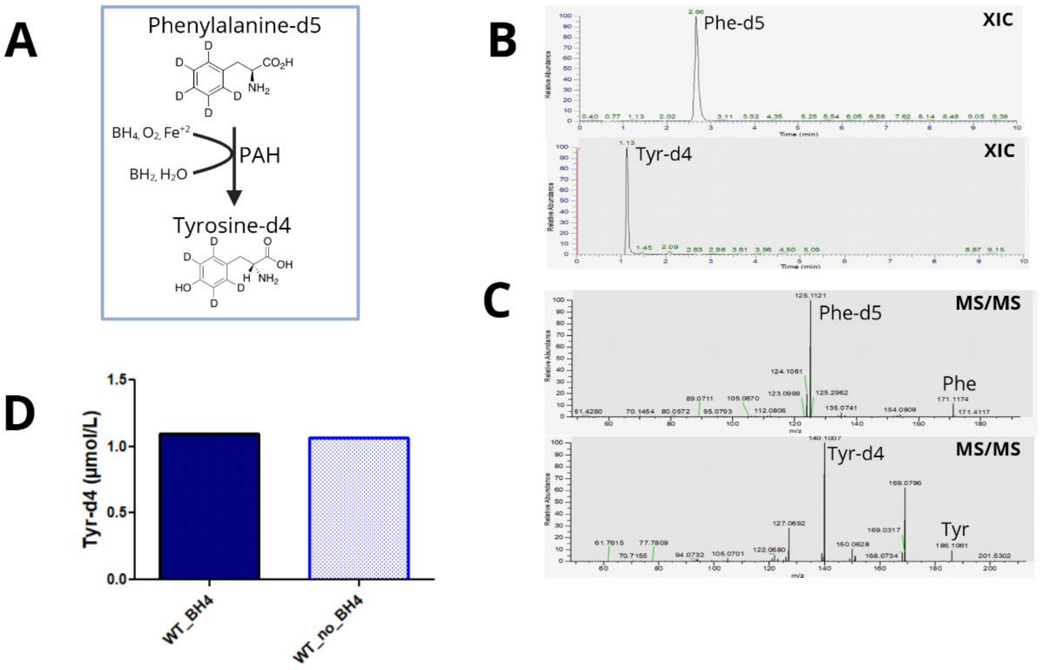


Figure 3.11:Development of an *in vitro* PAH enzymatic activity assay using LC-MS/MS.

(A) Schematic of phenylalanine-d5 conversion into tyrosine-d4 by PAH in the presence of BH4, oxygen, and iron. (B) Extracted ion chromatograms (XICs) of phenylalanine-d5 and tyrosine-d4 from standard sample spiked with protein extract. Phenylalanine-d5 extracted mass was 171.1176 m/z and tyrosine-d4 extracted mass was 186.1063 m/z. (C) Representative MS/MS spectra confirming the precursor and product ion transitions of phenylalanine-d5 and tyrosine-d4 from standard samples spiked with protein extract. Parent mass was 1171.1176 m/z and 186.1063 m/z respectively. (D) Example of an experimental reaction in which no Tyrosine-d4 was detected, indicating that further optimisation of the assay conditions is required. Assay optimisation, sample preparation, LC-MS/MS and data acquisition were performed by Dr Mark Graham and Dr Md Musfizur Hassan (CMRI, Biomedical Proteomics). M/z; mass-to-charge ratio.

`

3.3 Discussion

This study aimed to develop a mutant-agnostic gene editing approach to treat PKU by targeting the *Pah* locus with a dual CRISPR/Cas9 homology independent targeted integration (HITI) system delivered by AAV8. While we demonstrated successful targeting genomic integration at intron 1, generation of hybrid transcripts, and reporter gene expression, the approach did not achieve functional correction of hyperphenylalaninemia. The following discussion considers our results in the context of integration outcomes, hybrid transcript discordance despite other molecular analyses, GFP expression, and functional outcomes.

Integration outcomes

On-target integration at the Pah locus was confirmed at the genomic level by sequencing and quantification of junctions between donor DNA and intron 1. Correct junctions averaged 1.5-9.2% and reverse junctions 1.4-6.4% across GFP and PAH donor groups, with no significant difference observed between vector configurations or genetic cargo. To further characterise events at the cut site, PCR amplification was performed across the sgRNA cleavage site to assess the number of recoverable amplicons. Recovery of PCR products spanning the cut site was markedly reduced in GFP cohorts and trended downwards in PAH cohorts, consistent with targeted integration disrupting amplification. Strikingly, however, the combined proportion of correct junctions, reverse junctions and recoverable cut site PCR products did not account for all the alleles, leaving a "missing fraction". A second observation was the apparent discordance between correct genomic junction frequency and GFP reporter expression. For example, in the mouse with the highest proportion of GFP-positive hepatocytes by FACS (51.7%), only 3.8% correct and 2.3% reverse junctions were detected, with 70.5% recoverable cut-site PCR products. This mouse was found to have 18% indels, which at least by ICE, only saw small insertions and deletions, although this may be biased as a PCR approach.

These discrepancies likely reflect the limitations of this PCR-based assay, which has been configured to amplify the specific expected perfect HITI junctions. This “missing fraction” likely consists of complex integration outcomes such as the incorporation of concatermised AAV genomes or subgenomic rearranged fragments that may be functional would not be captured (45,47,233). Previous work from our group targeting the *Otc* locus using the same deep intronic HITI approach demonstrated such events by long-read sequencing (210). Consistent with this, others have shown that concatemer integration is a frequent outcome of rAAV editing and is mediated by the ITRs (233,234). These findings emphasise that PCR-based assays underestimate the true diversity of AAV-mediated editing outcomes and that long-read sequencing to quantify all functional outcomes will be essential in future studies.

Mechanistically, initial donor integration is thought to be equally probably in either orientation. HITI-donors include Cas9 recognition sites flanking the sequence for integration, allowing re-cleavage of reverse orientated inserts and thereby enriching for the correct orientation over time as long as the guide sites remain intact. In our study, a modest bias in favour of correct junctions was observed (1.3-1.6 correct:1 reverse). We did not observe a significant difference in the proportion of correct junctions between the GFP and PAH cohorts, despite the PAH cohort surviving for an additional four weeks. Although a great number of events, and a shift towards a higher correct-to-reverse ratio, might be expected over time when using HITI, this was not observed. Several explanations may account for this finding. First, by four weeks post-vector delivery, most integration events may have already occurred, including those that disrupt PAM sites, compassing both correct integrations and other events such as sub-genomic fragment and full AAV vector integration events. Alternatively, the relatively short difference in sampling time between GFP and PAH groups (4 weeks versus 8.5 weeks post-vector) may be insufficient to resolve temporal differences. Finally, the overall low recovery of junctions may limit the dynamic range needed to detect meaningful differences.

Transcript outcomes

We saw a significant difference in the amount of hybrid transcript between GFP or PAH donor sequence, with a range of 15.5-44.5% for dual treated GFP mice and 1.6-8.9% for dual treated PAH mice. Notably GFP transcripts were higher than PAH transcript despite comparable indel rates, junction frequencies, and vector copy numbers. For example, the PAH mice with the highest detected hybrid transcript (#213) has a higher recoverable correction junction as well as Cas9 VCN than 70% of the GFP treated cohort. This discordance suggests either an assay limitation for *PAH* or biology that differentially depresses *PAH* levels.

A technical explanation appears most likely. The dPCR for PAH transcript showed poor recovery when a known amount of purified hybrid PAH PCR product was spiked into the reaction, indicating suboptimal assay efficiency. While dPCR is typically less sensitive to amplification inefficiencies than qPCR, this may nonetheless account for the under-quantification of PAH transcripts. Biological reasons cannot be fully excluded. PAH transcript may lack stabilising elements or acquire destabilising ones (AU-rich elements), leading to reduced half-life (235). Taken together, the lower PAH transcript abundance likely reflects a combination of assay under-recovery of PAH hybrid isoforms rather than differences in editing frequency compared to the cohort receiving the donors containing the EGFP transgene. Future studies need to do perform further assay optimisation.

An interesting relationship was observed between GFP transcript abundance and proportion of GFP-positive hepatocytes by IHC (explored in more detail below). At lower levels of GFP expression (below 40% GFP positive by IHC), transcript levels correlated closely with the highest proportion of GFP-positive cells (1:1 ratio). However, in the two mice with matched data with the highest GFP%, (63% and 73% by IHC), transcript levels were lower (38-39%). By contrast, a third mouse with 43% IHC GFP positive had transcript levels of 45%. This apparent “shift” at higher expression levels may be explained by allelic editing. Since GFP positivity requires editing only one alle, cells become GFP positive regardless of whether one or both alleles are targeted. As more cells undergo editing, the probability increases that only one allele per cells is targeted, which could lower the overall ratio of transcript levels to GFP-positive cells.

GFP expression

Both the hybrid and universal GFP donor vectors resulted in robust GFP expression, detected by both immunohistochemistry and flow cytometry, with close correlation between the two methods. Analysis of GFP expression revealed three key features: a pericentral preference of GFP staining, a difference in GFP intensity between the hybrid and universal donor approaches, and a small (<1%) population of GFP-positive cells in donor-only cohorts.

The first feature was the enrichment of EGFP expression around the pericentral region, which could be due to two factors: differential expression of PAH expression across the liver influencing GFP expression, or differential transduction efficiency of the AAV8 capsid across the liver. Approximately half of all liver genes are expressed in a zonated manner (236), with amino acid catabolism generally enriched in periportal zones where oxygen and substrate concentrations are higher. Supporting this, a 2010 study cultured rat hepatocytes under different oxygen concentrations to mimic periportal or perivenous regions and found that lower oxygen significantly reduced PAH mRNA, protein, and enzyme activity. This indicates that PAH expression is positively regulated by oxygen levels and therefore expected to be higher in periportal regions (237). However, extensive studies using single-cell and spatial transcriptomics in both mouse (236) and human liver (238–241) documented many genes with strong zonation patterns, but PAH has not been among them. Furthermore, immunohistochemistry shows relatively low PAH expression throughout the liver, suggesting that PAH is not zonated in a meaningful way (119). Taken together, these findings indicate that zonated PAH expression is unlikely to account for the pericentral enrichment of GFP observed in our study.

By contrast, capsid-specific transduction preference is a well-documented phenomenon. Pseudo-serotyped rAAV2/8 vectors efficiently transduce murine hepatocytes, with most cells expressing a GFP cassette in previous AAV gene addition studies (163,215). Nevertheless, AAV8 capsid has been reported to transduce hepatocytes surrounding the central vein in mice (163,242,243) and dogs (244). This likely explains the higher density of GFP-positive hepatocytes observed in pericentral regions in our study.

Interestingly, the tropism of AAV appears to be species-specific. In cynomolgus and rhesus macaques, AAV preferentially transduced periportal hepatocytes rather than pericentral ones (244). Cabanes-Cruz and colleagues (2023) (245) proposed that this difference may reflect variations in AAVR composition and distribution as well as differences in extracellular matrix

`

proteoglycans. Furthermore, AAV8 and other naturally occurring capsids generally are poor at transducing human hepatocytes compared to modern, engineered capsids such as LK03 and SYD12 (129,246). Although AAV8 itself is unlikely to be selected for use in future clinical trials, these findings highlight the importance of considering capsid-specific and species-specific transduction preferences when developing AAV-based gene therapies.
The second feature was that the 2A donor configuration produced lower GFP signal than the splice donor configuration. This is evident in microscopy images (brighter staining observed in mice treated with the hybrid GFP approach) and in mean fluorescence intensity (MFI) by flow cytometry. This does not necessarily indicate that the P2A construct was less effective at genomic integration but rather that its reliance on ribosomal skipping yields lower per cell protein expression, potentially falling below the detection threshold in some cells, especially in those with less PAH promoter activity across the porto-central axis.

The third feature was that low levels of GFP expression were also detected in donor-only cohorts (<1%). As these animals did not receive Cas9 or sgRNA and had no recoverable junctions or hybrid transcript, the GFP expression observed did not originated from targeted insertion at the *Pah* locus. Instead, expression likely results from off-target events. A well-described mechanism is random integration of rAAV genomes or fragments at sites of DNA damage or high transcriptional activity, reported variably in the literature with some studies reporting 1-3% of vector genomes randomly integration into the genome (45,46,247) while others suggest true random integration frequency is much lower at 0.001-0.7% (47). This is likely dependent on quantity of rAAV genome entering the cells, if AAV genome has sequence homology with chromosomal sequence leading to targeted recombination at those loci and time following vector administration. Another possible explanation is episomal expression initiated from the ITRs, though this mechanism is unlikely to fully account for the

`

levels observed here, as ITR-derived promoter activity is generally weak and would be liver wide (248,249). As PAH IHC could not be assessed, a functional assay was pursued as a proxy for WT PAH expression.

Functional outcomes

The mutant *Pah*$^{enu2/enu2}$ mice treated with dual CRISPR/Cas9 and either PAH donor vector did not show any significant reduction in blood phenylalanine at 8 weeks post treatment. Two mice, however, had a modest decrease in blood phenylalanine over time (32.5 % and 49.4%). One mouse (#213) even exhibited a blunted response to dietary protein load, maintaining stable phenylalanine levels accompanied by an increase in blood tyrosine levels, which suggests that the additional dietary phenylalanine was successfully metabolised into tyrosine. This mouse also has the highest Cas9 and PAH donor vector copy number, the highest transcript detected (8.9%), the highest indels detected (16%) and the second highest correct junctions detected (6.6%).

These results, when considering the proportion of GFP% cells in GFP donor cohorts, are surprising. It has been suggested previously that there is a low threshold required to rescue hyperphenylalaninemia (Table 1.3) for which every dual treated GFP mouse assayed surpassed (IHC% range 18.27-72.7%). However, this assumes that the PAH protein is both present in equal quantities and assumes that any PAH present is functional. These results, when considering the PAH transcription data, are not surprising. Although it seems likely that a technical reason can explain this discrepancy, taken on face value, the low levels of PAH transcript achieved (range 1.6-8.9%) do not reach the threshold that has been seen in previous studies in this animal model (transcript 30-35% required) (Table 1.3). However, as previously discussed, *Pah* transcript levels may be underestimated. Looking at the GFP cohort, 4 out of the 10 total dual Cas9 GFP donor mice had >35% Hybrid GFP transcript expressed. These four mice had the highest GFP protein expression by IHC, the highest indels frequency (11-18%), highest Cas9 VCN (~23-31) and variable correct junctions (~4-9%). Among the PAH-treated cohort, mouse #213 exhibited the highest levels of editing, with comparable indels (16%), VCN (~25) and correct junctions (6.5%). If indeed the *Pah* hybrid transcript assay is not working well, and assuming GFP cohort can predict PAH cohort performance, this mouse may have been on the cusp of the therapeutic threshold. Regardless, the current dual-vector strategy is insufficient to reduce blood phenylalanine levels to a therapeutic range.

The *in vitro* PAH enzymatic assay is a standard method for assessing therapeutic efficacy in PKU studies. This is particularly important in experiments using the *Pah*$^{enu2/enu2}$ mouse model which are employed in most preclinical studies. In this model, conventional immunohistochemistry is not informative, as available PAH antibodies cannot distinguish between wildtype and mutant protein. In this study, we attempted to establish an assay in which liver lysate containing PAH catalysed labelled phe-d5 to into labelled tyrosine-d4, which could be measured as output of enzymatic activity. Samples were separated by liquid chromatography and analysed by tandem mass spectrometry, following methods reported in the literature. LC-MS/MS conditions were successfully optimised to detect both isotopically labelled substrate and product. However, the assay has not yet been fully optimised to reliably measure PAH activity in liver lysates. One potentially technical limitation was the use of aged catalase. Catalase protects BH4 from auto-oxidation. In its absence, hydrogen peroxide accumulates, which can oxidise BH4, Fe and PAH itself leading to a loss of enzymatic activity. While *in vitro* PAH assay has been widely described in the literature, there is considerable variation between protocols, and subtle difference in cofactor may have contributed to reduced assay sensitivity in this case. Despite laying important technical

`

groundwork, further optimisation will be required before the assay can provide reliable quantitative measures of PAH activity. Moreover, given the modest amounts of hybrid PAH transcript detected in this study and a non-significant reduction in blood phenylalanine levels observed relative to those reported in the literature, it is possible that differences in enzymatic activity would have been challenging to detect in our experimental groups even with a fully optimised assay.

An interesting observation was the reversion of coat colour from light to dark brown observed exclusively in dual CRISPR/Cas9 PAH donor treated mice. Coat colour reversion is commonly observed in preclinical gene therapy studies using the Pah$^{enu2/enu2}$ mouse, even in cases where blood phenylalanine is not substantially reduced. For example, in a similar study to ours, Richards and colleagues (119) introduced a therapeutic cassette using a dual CRISPR/Cas9 HDR approach and observed coat colour reversion at blood phenylalanine ranging between 700-1200 µM, comparable to those observed in this study. Taken together, these observations suggest a partial increase in PAH activity sufficient to increase tyrosine availability enough to restore melanin production in melanocytes. It is important to remember however, that it is not the lack of systemic tyrosine that is responsible for the pathology seen in PKU, but rather the reduced availability of tyrosine and tryptophan within the brain due to competitive inhibition by elevated phenylalanine. Melanocyte and neurons utilise different aromatic amino acid transporters (250,251) which may be less prone to competitive inhibition of phenylalanine.

Experimental design comments

An important consideration for achieving integration efficiency is both the total vector dose and the ratio of CRISPR/Cas9 to donor vector. The vector doses used in this study were selected based on previous data by our group, which identified optimal dose and dose ratios for achieving targeted integration using this HITI approach using similar editing reagents (213). In that work, variable doses and ratios of CRISPR/Cas9 and GFP donor vector were tested for GFP integration efficiency, and it was found a 2:1 ratio of CRISPR/Cas9 to Donor vector ($1 \times 10^{12}:5 \times 10^{11}$) was found to be optimal. This finding suggests that DSB generation may be the limiting factor, either due to CRISPR/Cas9 protein availability or sgRNA concentration. Careful titration of the combined vector dose is important to minimise the likelihood of AAV vector toxicity, while also considering the potential cost of future therapeutic production. Future work could explore strategies to improve integration efficiency, for example by incorporating the sgRNA sequence into the donor vector.

3.3.13 Conclusion

In summary, these findings demonstrate that a dual-vector HITI system can be used to effectively edit the Pah locus in the hepatocytes *in vivo*. Donor DNA was shown to integrate into *Pah* intron 1 at the genomic level, and result in RNA transcription. Protein expression resulting from targeted integration (up to 72% by IHC) was also observed after dual vector GFP treatment. These molecular and immunofluorescence data showed that the developed HITI system can lead to high levels of targeted integrations, albeit this was insufficient to achieve a therapeutic benefit in this case and therefore requires optimisation. Strategies such as alternative CRISPR delivery, improved guide selection, and careful dose escalation may collectively enhance both the efficacy and safety of this approach in subsequent preclinical investigations.

4. Discussion, conclusions and future perspectives

4.1 General discussion

This study aimed to develop a dual-AAV CRISPR/Cas9 gene editing approach to integrate a functional PAH cDNA sequence via HITI to provide a mutation-agnostic treatment for PKU by targeting *Pah* intron 1. To achieve this, we developed two general HITI approaches: a “hybrid” approach to treat coding mutations from exon 2-13, and a “universal” approach capable of addressing all pathogenetic coding mutations. Both approaches are advantageous when considering future clinical application because safety profiling will only have to be performed once thereby lessening testing as opposed to using unique editing reagents on a mutation-to-mutation basis and ease of manufacturing.

Using this approach, we achieved high levels of GFP expression, detectable hybrid transcripts, and confirmed targeted integration at the target site, intron 1 of *Pah*. However, dual CRISPR/Cas9 PAH donor approach did not result in any meaningful reduction in blood phenylalanine levels in Pahenu$^{2/enu2}$ mice. This indicates that the integration efficiency was insufficient to increase PAH activity to a therapeutic threshold, and that further optimisation of the donor construct may be required. Additionally, we were unable to reliably quantify Pah hybrid transcript or PAH enzymatic activity *in vitro*, which limited our ability to compare the results seen using GFP reporter constructs to that of PAH therapeutic donors. Interestingly, coat colour reversion from light to dark fur was observed in dual CRISPR/Cas9 PAH donor treated mice, suggesting a partial increase in PAH activity sufficient to restore melanin production in the fur.

Our findings were broadly consistent with previous gene editing approaches for PKU. In our GFP reporter constructs, we observed a relatively high proportion of hybrid transcripts (15-

`

44%). However, this was not recapitulated in the PAH donor treated mice, likely reflecting experimental limitations. Even so, these levels may not have been close to the threshold required for therapeutic benefit.

The study most comparable to ours is Richard et al.(119), which to date is the only other CIRSPR/Cas9 based approach for PKU, though it employed a HDR repair strategy. They achieved higher levels of allelic correct on average to us (~10%) and reported transcript levels within a similar range (2-24%) Importantly, this translated into approximately 10% wildtype PAH enzymatic activity, a level considered sufficient for metabolic correction according to some cell transplantation studies.

Among preclinical studies that successful reduced blood phenylalanine the lowest reported requirements were described in Villiger et al., (193) using a cytidine base editor. In their model, an average of 18% allele correction and at least 24% wildtype PAH mRNA was sufficient to normalise phenylalanine levels. Together, these studies provide useful benchmarks, suggesting that while our observed transcript results may have been encouraging, they were likely still borderline if recapitulated in the PAH donor treated mice.

4.2 Limitations and Future directions

This study had several important limitations that must be considered when interpreting the results. Technical issues, particularly with PAH transcript quantification, may have contributed to variability in the data. Furthermore, functional rescue was not achieved, likely reflecting the combined challenges of low integration frequency, limited expression, and constraints of the dual-vector approach. Inherently, dual-vector systems are restricted by co-transduction efficiency, and it is likely that higher doses, especially of the Cas9 vector, would

`

be required to improve the probability to targeted integration events and to achieve meaningful reductions in blood phenylalanine. An expanded dose-escalation study would determine if therapeutic levels could be achieved using this editing approach which could then be further optimised.

Future studies should focus on refining vector design and the exploration of safety considerations to improve the therapeutic potential of this approach. These include moving away from the HITI strategy and towards the implementation of bidirectional donor cassettes, alternative delivery of CRISPR/Cas9 using LNP, guide optimisation and vector dosing considerations. Finally, as previously mentioned, high vector doses could improve transduction efficiency and integration rates. However, dose escalation must be carefully balanced against toxicity, because AAV-induced hepatotoxicity and immune responses are known to correlated with the number of vector genomes delivery to the nucleus (252).
The donor vectors in this study were designed with flanking sgRNA sites to enable HITI-mediate integration, favouring the correct insert orientation enabling transgene expression. An alternative strategy would be the development of bidirectional donor cassettes, in which the full PAH coding sequence is included in both forward and reverse orientations without the need for flanking guide sites. Such designs would ensure functional expression regardless of the insert orientation and may also harness concatemer integration events which are facilitated by the AAV ITRs (233,234), which were lost in the current HITI design. This bidirectional cassette strategy is currently under development for a range of genetic diseases, such as Intellia's preclinical work on a dual AAV LNP therapy for haemophilia, which looks to insert human K9 into safe harbour albumin (253).

`

Another strategy to increase editing efficiency would be to conduct a more comprehensive guide RNA screen within the PAH locus to identify sites with improved cutting and integration frequency. This could be done with SaCas9, which was used for this study and is small enough for AAV packaging, or with SpCas9 if opting for LNP-delivery, which is more extensively studied and has broader sgRNA flexibility. While SpCas9 is generally too large to fit into a single AAV vector, this is not a limitation when delivering RNA in an LNP, as LNP have no physical size restriction of payload (254). SpCas9 has wider range of in silico to facilitate guide optimisation, and studies have optimised sgRNA to be highly modified for better performance *in vivo* (255). However, the less stringent sgRNA constraints of SpCas9 carry an increased risk of off-target cutting, which must be carefully evaluated.

Another promising avenue to optimise genome editing for both safety and efficacy is to deliver the CRISPR/Cas9 machinery as RNA using lipid nanoparticles (LNPs). LNP-mediated delivery offers several advantages: (i) it avoids potential integration of Cas9 vector genomes, thereby increasing safety and in knock in studies, reducing competition at the cut site, (ii) it lowers the cumulative AAV dose when compared with rAAV-only systems, decreasing vector-related toxicity, (iii) transient expression of Cas9 limits nuclease exposure and may reduce the risk of off-target activity and (iv) it enables the use of larger Cas9, such as SpCas9, which cannot be packaged into a single AAV vector but remain the best characterised for genome editing (52,52,256,257). While transient nature of Cas9 could limit efficacy, it has been shown that repeated doses are LNP-Cas9 are well tolerated. In a preclinical study looked to induce exon skipping in the dystrophin gene to treat DMD, repeated intramuscular injections were repeated up to six times, and saw low immunogenicity (258). Furthermore, there are real efforts going into the improving the longevity of delivered mRNA by using high modified guide RNAs and circular RNA (51,259).

To date, most LNP-CRISPR/Cas9 approaches have focused on allele disruption as a therapeutic mechanism. The only approved LNP-CRISPR/Cas9 gene editing product to date, Casgevy, is used to modify hematopoietic stem cells *ex vivo* to disrupt the erythroid-specific enhancer of BCL11A. This results in the de-repression of foetal haemoglobin, which substitutes for defective adult haemoglobin and effectively treats transfusion-dependent B-thalassemia and sickle cell disease (14,260). Similar approaches are being developed *in vivo*, with the strongest evidence in clinical and preclinical studies for diseases such as transthyretin amyloidosis (NTLA-2001), hereditary angioedema (NTLA-2002), haemophilia, hypercholesterolemia and glioblastoma (52–54,261). Collectively, these examples demonstrate that LNP-Cas9 knockout therapy has already been reached in the clinic, leading to opportunities for its use in for gene knock in applications.

As LNPs cannot currently efficiently deliver DNA to the nucleus, a dual LNP CRISPR/Cas9 donor approach is not currently possible. As such, there has been surge in dual LNP-AAV studies where the editing machinery is delivered in the LNP transiently, while the template DNA is delivered in an AAV vector. This has been applied in depth to treat haemophilia A and B in Cheong Yeom laboratory, with back to back papers published in 2023 looking to both introduce either a human F8 or human F9 sequence into the safe-harbour *SerpcinC1* locus in mouse models of haemophilia using HDR (51,52). Both approaches blood coagulation activity and reduced haemophilic complications at low levels of knock-in efficiency (≤3% KI) without adverse effects. They did, however, still report AAV integration at the cut site in both studies. Preclinical studies are also being performed by Intellia, looking to integrate hK9 into albumin safe harbour, no published data to date (253). Other Dual LNP-AAV gene knock into studies have looked to treat hereditary tyrosinemia in mice (262) and *in vitro* models of

`

cystic fibrosis (263), although again with low rates of targeted integration (3.5-6%). These show promise for the use of dual LNP-CRISPR/Cas9 AAV donor approaches to treat PKU. Alongside CRISPR/Cas9 editing approaches, base and prime editing are being investigated to treat PKU and other liver metabolic disorders. There has been a surge of interest in treating preclinical mouse models that have corrected *Pah* mutations, leading to normalisation of blood phenylalanine levels (193–196). Moreover, in 2025, the first *in vivo* clinical application of base editing therapy was reported in a child with carbamoyl phosphate synthetase 1 (CPS1) deficiency (264). In this case, the authors rapidly generated and validated personalised reagents specific to the patient specific mutation (Q335X). The entire process, from sequencing the patient mutation to approval by FDA, took approximately 6 months. The infant, who receive increasing doses of the LNP-base editor began tolerating more protein in his diet and demonstrated metabolic resilience, underscoring the transformative potential of base editing therapy. Nonetheless, the authors themselves acknowledge the immense practical and regulatory challenges of developing bespoke reagents of each rare mutation. Base and prime editors, while powerful, current face considerable hurdles in manufacturing, regulation and equitable access. As such, while of great scientific interest, these approaches are not yet broadly translatable to the clinic, particularly for genetically heterogeneous diseases such as PKU.

4.3 Conclusions

In conclusion, this thesis established and evaluated two distinct AAV-CRISPR/Cas9 HITI donor strategies aimed at correcting most pathogenic variants underlying phenylketonuria (PKU). These studies demonstrated efficient genomic targeting, robust reporter expression, and production of hybrid transcript products, illustrating the feasibility of targeted integration at the *PAH* locus. However, blood phenylalanine concentrations were not reduced to

`

therapeutic levels, underscoring the persistent challenge of translating efficient editing events into sufficient metabolic correction *in vivo.*

These limitations nevertheless provide a clear roadmap for future work. Dual-modality approaches, in which LNPs deliver Cas9 machinery and AAVs provide donor templates, offer a promising route to increase efficiency, lower viral burden, and mitigate the risks of long-term nuclease expression. Incorporating bidirectional cassettes may also increase integration efficiency and achieve the enzyme activity required for phenotypic rescue.

More broadly, the work presented here contributes to the expanding evidence that CRISPR-based gene editing can be adapted to metabolic diseases where correction across diverse mutations is needed. The findings presented here highlight both the promise and the current limitations of genome editing technologies: demonstrating that efficient on-target editing and transgene expression are achievable, while emphasizing that delivery, regulation of expression, and safety remain critical barriers to clinical translation. With continued advances in delivery systems and cassette design, the strategies developed in this thesis represent important steps toward a universal, one-time gene editing therapy for PKU, and lay the foundation for extending such transformative approaches to a broader spectrum of inherited metabolic disorders.

5. References

1. Hillert A, Anikster Y, Belanger-Quintana A, Burlina A, Burton BK, Carducci C, et al. The Genetic Landscape and Epidemiology of Phenylketonuria. The American Journal of Human Genetics. 2020 Aug;107(2):234–50.

2. Ginn SL, Mandwie M, Alexander IE, Edelstein M, Abedi MR. Gene therapy clinical trials worldwide to 2023—an update. The Journal of Gene Medicine. 2024 Aug;26(8):e3721.

3. Naldini L. Ex vivo gene transfer and correction for cell-based therapies. Nat Rev Genet. 2011 May;12(5):301–15.

4. Bulcha JT, Wang Y, Ma H, Tai PWL, Gao G. Viral vector platforms within the gene therapy landscape. Sig Transduct Target Ther. 2021 Feb 8;6(1):53.

5. Wold W, Toth K. Adenovirus Vectors for Gene Therapy, Vaccination and Cancer Gene Therapy. CGT. 2014 Jan 31;13(6):421–33.

6. Khanal S, Ghimire P, Dhamoon A. The Repertoire of Adenovirus in Human Disease: The Innocuous to the Deadly. Biomedicines. 2018 Mar 7;6(1):30.

7. Mast TC, Kierstead L, Gupta SB, Nikas AA, Kallas EG, Novitsky V, et al. International epidemiology of human pre-existing adenovirus (Ad) type-5, type-6, type-26 and type-36 neutralizing antibodies: Correlates of high Ad5 titers and implications for potential HIV vaccine trials. Vaccine. 2010 Jan;28(4):950–7.

8. Shirley JL, De Jong YP, Terhorst C, Herzog RW. Immune Responses to Viral Gene Therapy Vectors. Molecular Therapy. 2020 Mar;28(3):709–22.

9. Raper SE, Chirmule N, Lee FS, Wivel NA, Bagg A, Gao G ping, et al. Fatal systemic inflammatory response syndrome in a ornithine transcarbamylase deficient patient following adenoviral gene transfer. Molecular Genetics and Metabolism. 2003 Sep;80(1–2):148–58.

10. Peng Z. Current Status of Gendicine in China: Recombinant Human Ad-p53 Agent for Treatment of Cancers. Human Gene Therapy. 2005 Sep;16(9):1016–27.

11. Naldini L, Blömer U, Gallay P, Ory D, Mulligan R, Gage FH, et al. In Vivo Gene Delivery and Stable Transduction of Nondividing Cells by a Lentiviral Vector. Science. 1996 Apr 12;272(5259):263–7.

12. Schröder ARW, Shinn P, Chen H, Berry C, Ecker JR, Bushman F. HIV-1 Integration in the Human Genome Favors Active Genes and Local Hotspots. Cell. 2002 Aug;110(4):521–9.

13. Maude SL, Frey N, Shaw PA, Aplenc R, Barrett DM, Bunin NJ, et al. Chimeric Antigen Receptor T Cells for Sustained Remissions in Leukemia. N Engl J Med. 2014 Oct 16;371(16):1507–17.

14. Frangoul H, Locatelli F, Sharma A, Bhatia M, Mapara M, Molinari L, et al. Exagamglogene Autotemcel for Severe Sickle Cell Disease. N Engl J Med. 2024 May 9;390(18):1649–62.

15. Eichler F, Duncan CN, Musolino PL, Lund TC, Gupta AO, De Oliveira S, et al. Lentiviral Gene Therapy for Cerebral Adrenoleukodystrophy. N Engl J Med. 2024 Oct 10;391(14):1302–12.

16. Duncan CN, Bledsoe JR, Grzywacz B, Beckman A, Bonner M, Eichler FS, et al. Hematologic Cancer after Gene Therapy for Cerebral Adrenoleukodystrophy. N Engl J Med. 2024 Oct 10;391(14):1287–301.

17. Montini E, Cesana D, Schmidt M, Sanvito F, Bartholomae CC, Ranzani M, et al. The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. J Clin Invest. 2009 Apr 1;119(4):964–75.

18. Ginn SL, Christina S, Alexander IE. Genome editing in the human liver: Progress and translational considerations. In: Progress in Molecular Biology and Translational Science [Internet]. Elsevier; 2021 [cited 2025 Aug 31]. p. 257–88. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1877117321000417

19. Hastie E, Samulski RJ. Adeno-Associated Virus at 50: A Golden Anniversary of Discovery, Research, and Gene Therapy Success—A Personal Perspective. Human Gene Therapy. 2015 May;26(5):257–65.

20. Philpott NJ, Giraud-Wali C, Dupuis C, Gomos J, Hamilton H, Berns KI, et al. Efficient Integration of Recombinant Adeno-Associated Virus DNA Vectors Requires a p5- *rep* Sequence in *cis*. J Virol. 2002 Jun;76(11):5411–21.

21. La Bella T, Imbeaud S, Peneau C, Mami I, Datta S, Bayard Q, et al. Adeno-associated virus in the liver: natural history and consequences in tumour development. Gut. 2020 Apr;69(4):737–47.

22. Nault JC, Datta S, Imbeaud S, Franconi A, Mallet M, Couchy G, et al. Recurrent AAV2-related insertional mutagenesis in human hepatocellular carcinomas. Nat Genet. 2015 Oct;47(10):1187–93.

23. Rose JA, Maizel JV, Inman JK, Shatkin AJ. Structural Proteins of Adenovirus-Associated Viruses. J Virol. 1971 Nov;8(5):766–70.

24. Meyer NL, Hu G, Davulcu O, Xie Q, Noble AJ, Yoshioka C, et al. Structure of the gene therapy vector, adeno-associated virus with its cell receptor, AAVR. eLife. 2019 May 22;8:e44707.

25. Srivastava A. In vivo tissue-tropism of adeno-associated viral vectors. Current Opinion in Virology. 2016 Dec;21:75–80.

26. Lopez-Gordo E, Chamberlain K, Riyad J, Kohlbrenner E, Weber T. Natural Adeno-Associated Virus Serotypes and Engineered Adeno-Associated Virus Capsid Variants: Tropism Differences and Mechanistic Insights. Viruses. 2024 Mar 12;16(3):442.

27. Samulski RJ, Berns KI, Tan M, Muzyczka N. Cloning of adeno-associated virus into pBR322: Rescue of intact virus from the recombinant plasmid in human cells.

28. Choi VW, McCarty DM, Samulski RJ. Host Cell DNA Repair Pathways in Adeno-Associated Viral Genome Processing. J Virol. 2006 Nov;80(21):10346–56.

29. Hanlon KS, Kleinstiver BP, Garcia SP, Zaborowski MP, Volak A, Spirig SE, et al. High levels of AAV vector integration into CRISPR-induced DNA breaks. Nat Commun. 2019 Sep 30;10(1):4439.

30. Ferrari FK, Samulski T, Shenk T, Samulski RJ. Second-strand synthesis is a rate-limiting step for efficient transduction by recombinant adeno-associated virus vectors. J Virol. 1996 May;70(5):3227–34.

31. McCarty D, Monahan P, Samulski R. Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis. Gene Ther. 2001 Aug 1;8(16):1248–54.

32. McCarty DM, Fu H, Monahan PE, Toulson CE, Naik P, Samulski RJ. Adeno-associated virus terminal repeat (TR) mutant generates self-complementary vectors to overcome the rate-limiting step to transduction in vivo. Gene Ther. 2003 Dec 1;10(26):2112–8.

33. Calcedo R, Vandenberghe LH, Gao G, Lin J, Wilson JM. Worldwide Epidemiology of Neutralizing Antibodies to Adeno-Associated Viruses. J INFECT DIS. 2009 Feb;199(3):381–90.

34. Mingozzi F, High KA. Immune responses to AAV vectors: overcoming barriers to successful gene therapy. Blood. 2013 Jul 4;122(1):23–36.

35. ASGCT and FDA Liaison Meeting [Internet]. 2021. Available from: https://www.asgct.org/uploads/files/general/News/FINAL-combined-slides-redacted-2021-ASGCT-FDA-Liaison-Meeting.pdf

36. Ronzitti G, Gross DA, Mingozzi F. Human Immune Responses to Adeno-Associated Virus (AAV) Vectors. Front Immunol. 2020 Apr 17;11:670.

37. Nathwani AC, Tuddenham EGD, Rangarajan S, Rosales C, McIntosh J, Linch DC, et al. Adenovirus-Associated Virus Vector–Mediated Gene Transfer in Hemophilia B. N Engl J Med. 2011 Dec 22;365(25):2357–65.

38. Kishimoto TK, Samulski RJ. Addressing high dose AAV toxicity – ‘one and done’ or ‘slower and lower’? Expert Opinion on Biological Therapy. 2022 Sep 2;22(9):1067–71.

39. Novartis Pharmacetuicals UK Ltd. Onasemnogene abeparvovec, ZOLGENSMA - Fatal Cases of Acute Liver Failure [Internet]. HPRA; 2023. Available from: https://assets.hpra.ie/data/docs/default-source/product-updates/dhpc/human-medicines/zolgensma-%28onasemnogene-abeparvovec%29-direct-healthcare-professional-communication-%28dhpc%29-february-2023.pdf?sfvrsn=ed6c6fd2_1&utm

40. Chand D, Mohr F, McMillan H, Tukov FF, Montgomery K, Kleyn A, et al. Hepatotoxicity following administration of onasemnogene abeparvovec (AVXS-101) for the treatment of spinal muscular atrophy. Journal of Hepatology. 2021 Mar;74(3):560–6.

41. Shieh PB, Kuntz NL, Dowling JJ, Müller-Felber W, Bönnemann CG, Seferian AM, et al. Safety and efficacy of gene replacement therapy for X-linked myotubular myopathy (ASPIRO): a multinational, open-label, dose-escalation trial. The Lancet Neurology. 2023 Dec;22(12):1125–39.

42. Buss N, Lanigan L, Zeller J, Cissell D, Metea M, Adams E, et al. Characterization of AAV-mediated dorsal root ganglionopathy. Molecular Therapy - Methods & Clinical Development. 2022 Mar;24:342–54.

43. Nguyen GN, Everett JK, Kafle S, Roche AM, Raymond HE, Leiby J, et al. A long-term study of AAV gene therapy in dogs with hemophilia A identifies clonal expansions of transduced liver cells. Nat Biotechnol. 2021 Jan;39(1):47–55.

44. Asokan A, Shen S. Redirecting AAV vectors to extrahepatic tissues. Mol Ther. 2023 Dec 6;31(12):3371–5.

45. Dalwadi DA, Calabria A, Tiyaboonchai A, Posey J, Naugler WE, Montini E, et al. AAV integration in human hepatocytes. Molecular Therapy. 2021 Oct;29(10):2898–909.

46. Ismail AM, Witt E, Bouwman T, Clark W, Yates B, Franco M, et al. The longitudinal kinetics of AAV5 vector integration profiles and evaluation of clonal expansion in mice. Molecular Therapy - Methods & Clinical Development. 2024 Sep;32(3):101294.

47. Greig JA, Martins KM, Breton C, Lamontagne RJ, Zhu Y, He Z, et al. Integrated vector genomes may contribute to long-term expression in primate liver after AAV administration. Nat Biotechnol. 2024 Aug;42(8):1232–42.

48. Schmidt M, Foster GR, Coppens M, Thomsen H, Dolmetsch R, Heijink L, et al. Molecular evaluation and vector integration analysis of HCC complicating AAV gene therapy for hemophilia B. Blood Advances. 2023 Sep 12;7(17):4966–9.

49. Wang J, Ding Y, Chong K, Cui M, Cao Z, Tang C, et al. Recent Advances in Lipid Nanoparticles and Their Safety Concerns for mRNA Delivery. Vaccines. 2024 Oct 8;12(10):1148.

50. Hou X, Zaks T, Langer R, Dong Y. Lipid nanoparticles for mRNA delivery. Nat Rev Mater. 2021 Aug 10;6(12):1078–94.

51. Lee JH, Han JP, Song DW, Lee GS, Choi BS, Kim M, et al. In vivo genome editing for hemophilia B therapy by the combination of rebalancing and therapeutic gene knockin using a viral and non-viral vector. Molecular Therapy - Nucleic Acids. 2023 Jun;32:161–72.

52. Han JP, Lee Y, Lee JH, Chung HY, Lee GS, Nam YR, et al. In vivo genome editing using 244-cis LNPs and low-dose AAV achieves therapeutic threshold in hemophilia A mice. Molecular Therapy Nucleic Acids. 2023 Dec;34:102050.

53. Gillmore JD, Gane E, Taubel J, Kao J, Fontana M, Maitland ML, et al. CRISPR-Cas9 In Vivo Gene Editing for Transthyretin Amyloidosis. N Engl J Med. 2021 Aug 5;385(6):493–502.

54. Longhurst HJ, Lindsay K, Petersen RS, Fijen LM, Gurugama P, Maag D, et al. CRISPR-Cas9 In Vivo Gene Editing of *KLKB1* for Hereditary Angioedema. N Engl J Med. 2024 Feb;390(5):432–41.

55. Kuntz E, Kuntz HD. Biochemistry and Functions of the Liver. In: Hepatology Principles and Practice. 2nd ed. Springer Berlin Heidelberg; 2006.

56. Elias H. A re-examination of the structure of the mammalian liver. II. The hepatic lobule and its relation to the vascular and biliary systems. Am J Anat. 1949 Nov;85(3):379–456.

57. Horn T, Henriksen JH, Christoffersen P. The sinusoidal lining cells in “normal” human liver. A scanning electron microscopic investigation. Liver. 1986 Apr;6(2):98–110.

58. Wisse E, De Zanger RB, Charels K, Van Der Smissen P, McCuskey RS. The Liver Sieve: Considerations Concerning the Structure and Function of Endothelial Fenestrae, the Sinusoidal Wall and the Space of Disse. Hepatology. 1985 Jul;5(4):683–92.

59. Jeanmonod R, Asuka E, Jeanmonod D. Inborn Errors of Metabolism. [Updated 2023 Jul 17]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK459183/. In.

60. Van Spronsen FJ, Blau N, Harding C, Burlina A, Longo N, Bosch AM. Phenylketonuria. Nat Rev Dis Primers. 2021 May 20;7(1):36.

61. Fölling A. Über Ausscheidung von Phenylbrenztraubensäure in den Harn als Stoffwechselanomalie in Verbindung mit Imbezillität. Hoppe-Seyler´s Zeitschrift für physiologische Chemie. 1934 Jan;227(1–4):169–81.

62. Jervis G. Phenylpyruvic oligophrenia deficiency of phenylalanine-oxidizing system. Proceedings of the Society for Experimental Biology and Medicine. 1953;82(3).

63. Bickel H, Gerrard J, Hickmans E. Influence of phenylalanine intake on phenylketonuria. The Lancet. 1953;262(6790).

64. Guthrie R, Susi A. A SIMPLE PHENYLALANINE METHOD FOR DETECTING PHENYLKETONURIA IN LARGE POPULATIONS OF NEWBORN INFANTS. Pediatrics. 1963;32(3):338–43.

65. Moat SJ, George RS, Carling RS. Use of Dried Blood Spot Specimens to Monitor Patients with Inherited Metabolic Disorders. IJNS. 2020 Mar 26;6(2):26.

`

66. Ac I. Australasian consensus guidelines for the management of phenylketonuria (PKU) throughout the lifespan.

67. MacDonald A, Van Wegberg AMJ, Ahring K, Beblo S, Bélanger-Quintana A, Burlina A, et al. PKU dietary handbook to accompany PKU guidelines. Orphanet J Rare Dis. 2020 Dec;15(1):171.

68. Pinto A, Ilgaz F, Evans S, Van Dam E, Rocha JC, Karabulut E, et al. Phenylalanine Tolerance over Time in Phenylketonuria: A Systematic Review and Meta-Analysis. Nutrients. 2023 Aug 8;15(16):3506.

69. Owada M, Aoki K, Kitagawa T. Taste preferences and feeding behaviour in children with phenylketonuria on a semisynthetic diet. Eur J Pediatr. 2000 Oct 4;159(11):846–50.

70. Hafid NA, Christodoulou J. Phenylketonuria: a review of current and future treatments.

71. De Giorgi A, Nardecchia F, Romani C, Leuzzi V. Metabolic control and clinical outcome in adolescents with phenylketonuria. Molecular Genetics and Metabolism. 2023 Nov;140(3):107684.

72. Waisbren SE, Noel K, Fahrbach K, Cella C, Frame D, Dorenbaum A, et al. Phenylalanine blood levels and clinical outcomes in phenylketonuria: A systematic literature review and meta-analysis. Molecular Genetics and Metabolism. 2007 Sep;92(1–2):63–70.

73. Gundorova P, Yousefi B, Woidy M, Rose-Heine MS, Khatri R, Kasten V, et al. Personalized Genotype-Based Approach for Treatment of Phenylketonuria. J of Inher Metab Disea. 2025 Sep;48(5):e70067.

74. Schuck PF, Malgarin F, Cararo JH, Cardoso F, Streck EL, Ferreira GC. Phenylketonuria Pathophysiology: on the Role of Metabolic Alterations. Aging and disease. 2015;6(5):390.

75. Konecki DS. Structural Characterization of the 5' Regions of the Human Phenylalanine Hydroxylase Genet.

76. Gorissen SH, Trommelen J, Kouw IW, Kouw IW, Pennings B, Pennings B, et al. Protein Type, Protein Dose, and Age Modulate Dietary Protein Digestion and Phenylalanine Absorption Kinetics and Plasma Phenylalanine Availability in Humans. The Journal of Nutrition. 2020 Aug;150(8):2041–50.

77. harissa A. Dyer, Ady Kendler, Thomas Philibotte, Paula Gardiner, Jonathan Cruz, Harvey L. Levy,. Evidence for Central Nervous System Glial Cell Plasticity in Phenylketonuria.

78. McKEAN CM. THE EFFECTS OF HIGH PHENYLALANINE CONCENTRATIONS ON SEROTONIN AND CATECHOLAMINE METABOLISM IN THE HUMAN BRAIN.

79. Lou HC, Lykkelund C, Gerdes A -M., Udesen H, Bruhn P. Increased Vigilance and Dopamine Synthesis by Large Doses of Tyrosine or Phenylalanine Restriction in Phenylketonuria. Acta Paediatrica. 1987 Jul;76(4):560–5.

80. Pilotto A, Zipser CM, Leks E, Haas D, Gramer G, Freisinger P, et al. Phenylalanine Effects on Brain Function in Adult Phenylketonuria. Neurology [Internet]. 2021 Jan 19 [cited 2025 Aug 31];96(3). Available from: https://www.neurology.org/doi/10.1212/WNL.0000000000011088

81. González MJ, Gassió R, Artuch R, Campistol J. Impaired Neurotransmission in Early-treated Phenylketonuria Patients. Seminars in Pediatric Neurology. 2016 Nov;23(4):332–40.

82. Hoeksma M, Reijngoud DJ, Pruim J, De Valk HW, Paans AMJ, Van Spronsen FJ. Phenylketonuria: High plasma phenylalanine decreases cerebral protein synthesis. Molecular Genetics and Metabolism. 2009 Apr;96(4):177–82.

83. Imperlini E, Orrù S, Corbo C, Daniele A, Salvatore F. Altered brain protein expression profiles are associated with molecular neurological dysfunction in the PKU mouse model. Journal of Neurochemistry. 2014 Jun;129(6):1002–12.

84. Manek R, Zhang YV, Berthelette P, Hossain M, Cornell CS, Gans J, et al. Blood phenylalanine reduction reverses gene expression changes observed in a mouse model of phenylketonuria. Sci Rep. 2021 Nov 24;11(1):22886.

85. Adler-Abramovich L, Vaks L, Carny O, Trudler D, Magno A, Caflisch A, et al. Phenylalanine assembly into toxic fibrils suggests amyloid etiology in phenylketonuria. Nat Chem Biol. 2012 Aug;8(8):701–6.

86. Aubi O, Prestegård KS, Jung-KC K, Shi TJS, Ying M, Grindheim AK, et al. The Pah-R261Q mouse reveals oxidative stress associated with amyloid-like hepatic aggregation of mutant phenylalanine hydroxylase. Nat Commun. 2021 Apr 6;12(1):2073.

87. Li Y, Tan Z, Zhang Y, Zhang Z, Hu Q, Liang K, et al. A noncoding RNA modulator potentiates phenylalanine metabolism in mice. Science. 2021 Aug 6;373(6555):662–73.

88. Himmelreich N, Ramón-Maiques S, Navarrete R, Castejon-Fernandez N, Garbade SF, Martinez A, et al. Significance of utilizing in silico structural analysis and phenotypic data to characterize phenylalanine hydroxylase variants: A PAH landscape. Molecular Genetics and Metabolism. 2024 Jul;142(3):108514.

89. Wettstein S, Underhaug J, Perez B, Marsden BD, Yue WW, Martinez A, et al. Linking genotypes database with locus-specific database and genotype–phenotype correlation in phenylketonuria. Eur J Hum Genet. 2015 Mar;23(3):302–9.

90. Levy HL, Milanowski A, Chakrapani A, Cleary M, Lee P, Trefz FK, et al. Efficacy of sapropterin dihydrochloride (tetrahydrobiopterin, 6R-BH4) for reduction of phenylalanine concentration in patients with phenylketonuria: a phase III randomised placebo-controlled study. The Lancet. 2007 Aug;370(9586):504–10.

91. Zurflüh MR, Zschocke J, Lindner M, Feillet F, Chery C, Burlina A, et al. Molecular genetics of tetrahydrobiopterin-responsive phenylalanine hydroxylase deficiency. Hum Mutat. 2008 Jan;29(1):167–75.

92. Trefz FK, Burton BK, Longo N, Casanova MMP, Gruskin DJ, Dorenbaum A, et al. Efficacy of Sapropterin Dihydrochloride in Increasing Phenylalanine Tolerance in Children with Phenylketonuria: A Phase III, Randomized, Double-Blind, Placebo-Controlled Study. The Journal of Pediatrics. 2009 May;154(5):700-707.e1.

93. Burton BK, Nowacka M, Hennermann JB, Lipson M, Grange DK, Chakrapani A, et al. Safety of extended treatment with sapropterin dihydrochloride in patients with phenylketonuria: Results of a phase 3b study. Molecular Genetics and Metabolism. 2011 Aug;103(4):315–22.

94. Pharmacoeconomic Review Report: Sapropterin dihydrochloride (Kuvan) [Internet]. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2017. Available from: https://www.ncbi.nlm.nih.gov/books/NBK533821/

95. BIOPKU. 2014. Available from: https://www.biopku.org/home/pah.asp

96. Himmelreich N, Ramón-Maiques S, Navarrete R, Castejon-Fernandez N, Garbade SF, Martinez A, et al. Significance of utilizing in silico structural analysis and phenotypic data to characterize phenylalanine hydroxylase variants: A PAH landscape. Molecular Genetics and Metabolism. 2024 Jul;142(3):108514.

97. BIOPKU. Frequency of variants in the phenylalanine hydroxylase gene (PAH) [Internet]. 2024. Available from: https://www.biopku.org/home/pah.asp

98. Longo N, Harding CO, Burton BK, Grange DK, Vockley J, Wasserstein M, et al. Single-dose, subcutaneous recombinant phenylalanine ammonia lyase conjugated with polyethylene glycol in adult patients with phenylketonuria: an open-label, multicentre, phase 1 dose-escalation trial. The Lancet. 2014 Jul;384(9937):37–44.

99. Harding CO, Amato RS, Stuy M, Longo N, Burton BK, Posner J, et al. Pegvaliase for the treatment of phenylketonuria: A pivotal, double-blind randomized discontinuation Phase 3 clinical trial. Molecular Genetics and Metabolism. 2018 May;124(1):20–6.

100. Thomas J, Levy H, Amato S, Vockley J, Zori R, Dimmock D, et al. Pegvaliase for the treatment of phenylketonuria: Results of a long-term phase 3 clinical trial program (PRISM). Molecular Genetics and Metabolism. 2018 May;124(1):27–38.

101. Gupta S, Lau K, Harding CO, Shepherd G, Boyer R, Atkinson JP, et al. Association of immune response with efficacy and safety outcomes in adults with phenylketonuria administered pegvaliase in phase 3 clinical trials. EBioMedicine. 2018 Nov;37:366–73.

102. Hydery T, Coppenrath VA. A Comprehensive Review of Pegvaliase, an Enzyme Substitution Therapy for the Treatment of Phenylketonuria. Drug Target□Insights. 2019 Jan;13:117739281985708.

103. Vashishat A, Patel P, Das Gupta G, Das Kurmi B. Alternatives of Animal Models for Biomedical Research: a Comprehensive Review of Modern Approaches. Stem Cell Rev and Rep. 2024 May;20(4):881–99.

104. Westhaus A, Cabanes-Creus M, Dilworth KL, Zhu E, Salas Gómez D, Navarro RG, et al. Assessment of Pre-Clinical Liver Models Based on Their Ability to Predict the Liver-Tropism of Adeno-Associated Virus Vectors. Human Gene Therapy. 2023 Apr 1;34(7–8):273–88.

105. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021 Aug 26;596(7873):583–9.

106. Charbonneau MR, Denney WS, Horvath NG, Cantarella P, Castillo MJ, Puurunen MK, et al. Development of a mechanistic model to predict synthetic biotic activity in healthy volunteers and patients with phenylketonuria. Commun Biol. 2021 Jul 22;4(1):898.

107. Gjetting T, Romstad A, Haavik J, Knappskog PM, Acosta AX, Silva WA, et al. A Phenylalanine Hydroxylase Amino Acid Polymorphism with Implications for Molecular Diagnostics. Molecular Genetics and Metabolism. 2001 Jul;73(3):280–4.

108. Wang GA, Gu P, Kaufman S. Mutagenesis of the regulatory domain of phenylalanine hydroxylase. Proc Natl Acad Sci USA. 2001 Feb 13;98(4):1537–42.

109. Pey AL, Desviat LR, Gámez A, Ugarte M, Pérez B. Phenylketonuria: Genotype-phenotype correlations based on expression analysis of structural and functional mutations in *PAH*: STRUCTURAL/FUNCTIONAL PAH MUTATIONS. Hum Mutat. 2003 Apr;21(4):370–8.

110. Erlandsen H, Pey AL, Gámez A, Pérez B, Desviat LR, Aguado C, et al. Correction of kinetic and stability defects by tetrahydrobiopterin in phenylketonuria patients with certain phenylalanine hydroxylase mutations. Proc Natl Acad Sci USA. 2004 Nov 30;101(48):16903–8.

111. Bjørgo E, Knappskog PM, Martinez A, Stevens RC, Flatmark T. Partial characterization and three-dimensional-structural localization of eight mutations in exon 7 of the human phenylalanine hydroxylase gene associated with phenylketonuria. European Journal of Biochemistry. 1998 Oct;257(1):1–10.

112. Trunzo R, Santacroce R, Shen N, Jung-Klawitter S, Leccese A, De Girolamo G, et al. In vitro residual activity of phenylalanine hydroxylase variants and correlation with metabolic phenotypes in PKU. Gene. 2016 Dec;594(1):138–43.

113. Zhang X, Ye J, Shen N, Tao Y, Han L, Qiu W, et al. In vitro residual activities in 20 variants of phenylalanine hydroxylase and genotype-phenotype correlation in phenylketonuria patients. Gene. 2019 Jul;707:239–45.

114. Song J, Bang S, Choi N, Kim HN. Brain organoid-on-a-chip: A next-generation human brain avatar for recapitulating human brain physiology and pathology. Biomicrofluidics. 2022 Dec 1;16(6):061301.

115. Osonoi S, Takebe T. Organoid-guided precision hepatology for metabolic liver disease. Journal of Hepatology. 2024 May;80(5):805–21.

116. Shedlovsky A, McDonald JD, Symula D, Dove WF. Mouse models of human phenylketonuria. Genetics. 1993 Aug 1;134(4):1205–10.

117. Cho S, McDonald JD. Effect of Maternal Blood Phenylalanine Level on Mouse Maternal Phenylketonuria Offspring. Molecular Genetics and Metabolism. 2001 Dec;74(4):420–5.

118. Waters PJ, Scriver CR, Parniak MA. Homomeric and Heteromeric Interactions between Wild-Type and Mutant Phenylalanine Hydroxylase Subunits: Evaluation of Two-Hybrid Approaches for Functional Analysis of Mutations Causing Hyperphenylalaninemia. Molecular Genetics and Metabolism. 2001 Jul;73(3):230–8.

119. Richards DY, Winn SR, Dudley S, Nygaard S, Mighell TL, Grompe M, et al. AAV-Mediated CRISPR/Cas9 Gene Editing in Murine Phenylketonuria. Molecular Therapy - Methods & Clinical Development. 2020 Jun;17:234–45.

120. Singh K, Cornell CS, Jackson R, Kabiri M, Phipps M, Desai M, et al. CRISPR/Cas9 generated knockout mice lacking phenylalanine hydroxylase protein as a novel preclinical model for human phenylketonuria. Sci Rep. 2021 Mar 31;11(1):7254.

121. Martínez-Pizarro A, Picó S, López-Márquez A, Rodriguez-López C, Montalvo E, Alvarez M, et al. PAH DEFICIENT PATHOLOGY IN HUMANIZED c.1066-11G>A PHENYLKETONURIA MICE [Internet]. Pathology; 2023 [cited 2025 Aug 31]. Available from: http://biorxiv.org/lookup/doi/10.1101/2023.11.03.565447

122. Zhang C, Yan Y, Zhou B, Gao H, Jin X, Hui L, et al. Phenotypic study of humanized mice carrying the PAH deep intronic variant c.1199+502A>T. Orphanet J Rare Dis. 2025 May 26;20(1):254.

123. Koppes EA, Redel BK, Johnson MA, Skvorak KJ, Ghaloul-Gonzalez L, Yates ME, et al. A porcine model of phenylketonuria generated by CRISPR/Cas9 genome editing. JCI Insight. 2020 Oct 15;5(20):e141523.

124. Kaiser RA, Weber ND, Trigueros-Motos L, Allen KL, Martinez M, Cao W, et al. Use of an adeno-associated virus serotype Anc80 to provide durable cure of phenylketonuria in a mouse model. J of Inher Metab Disea. 2021 Nov;44(6):1369–81.

125. Challis RC, Ravindra Kumar S, Chen X, Goertsen D, Coughlin GM, Hori AM, et al. Adeno-Associated Virus Toolkit to Target Diverse Brain Cells. Annu Rev Neurosci. 2022 Jul 8;45(1):447–69.

126. Tricot T, Verfaillie CM, Kumar M. Current Status and Challenges of Human Induced Pluripotent Stem Cell-Derived Liver Models in Drug Discovery. Cells. 2022 Jan 27;11(3):442.

127. Corbett JL, Duncan SA. iPSC-Derived Hepatocytes as a Platform for Disease Modeling and Drug Discovery. Front Med. 2019 Nov 15;6:265.

128. Azuma H, Paulk N, Ranade A, Dorrell C, Al-Dhalimy M, Ellis E, et al. Robust expansion of human hepatocytes in Fah−/−/Rag2−/−/Il2rg−/− mice. Nat Biotechnol. 2007 Aug;25(8):903–10.

129. Cabanes-Creus M, Liao SHY, Gale Navarro R, Knight M, Nazareth D, Lau NS, et al. Harnessing whole human liver ex situ normothermic perfusion for preclinical AAV vector evaluation. Nat Commun. 2024 Mar 14;15(1):1876.

130. Kylies J, Brunne B, Rune GM. A culture model for the assessment of phenylalanine neurotoxicity in phenylketonuria. In vitro models. 2022 Jan 27;1(1):103–14.

131. Bruinenberg VM, Van Der Goot E, Van Vliet D, De Groot MJ, Mazzola PN, Heiner-Fokkema MR, et al. The Behavioral Consequence of Phenylketonuria in Mice Depends on the Genetic Background. Front Behav Neurosci [Internet]. 2016 Dec 20 [cited 2025 Aug 31];10. Available from: http://journal.frontiersin.org/article/10.3389/fnbeh.2016.00233/full

132. Haefele MJ, White G, McDonald JD. Characterization of the Mouse Phenylalanine Hydroxylase Mutation Pahenu3. Molecular Genetics and Metabolism. 2001 Jan;72(1):27–30.

133. Brooks DL, Carrasco MJ, Qu P, Peranteau WH, Ahrens-Nicklas RC, Musunuru K, et al. Rapid and definitive treatment of phenylketonuria in variant-humanized mice with corrective editing. Nat Commun. 2023 Jun 10;14(1):3451.

134. Kopple JD. Phenylalanine and Tyrosine Metabolism in Chronic Kidney Failure. The Journal of Nutrition. 2007 Jun;137(6):1586S-1590S.

135. Isabella VM, Ha BN, Castillo MJ, Lubkowicz DJ, Rowe SE, Millet YA, et al. Development of a synthetic live bacterial therapeutic for the human metabolic disease phenylketonuria. Nat Biotechnol. 2018 Sep;36(9):857–64.

136. Puurunen MK, Vockley J, Searle SL, Sacharow SJ, Phillips JA, Denney WS, et al. Safety and pharmacodynamics of an engineered E. coli Nissle for the treatment of phenylketonuria: a first-in-human phase 1/2a study. Nat Metab. 2021 Jul 22;3(8):1125–32.

137. Bratkovic D, Margvelashvili L, Tchan MC, Nisbet J, Smith N. PTC923 (sepiapterin) lowers elevated blood phenylalanine in subjects with phenylketonuria: a phase 2 randomized, multi-center, three-period crossover, open-label, active controlled, all-comers study. Metabolism. 2022 Mar;128:155116.

138. Smith N, Longo N, Levert K, Hyland K, Blau N. Exploratory study of the effect of one week of orally administered CNSA-001 (sepiapterin) on CNS levels of tetrahydrobiopterin, dihydrobiopterin and monoamine neurotransmitter metabolites in healthy volunteers. Molecular Genetics and Metabolism Reports. 2019 Dec;21:100500.

139. Wobst HJ, Viader A, Muncipinto G, Hollibaugh R, Van Kalken D, Burkhart CT, et al. SLC6A19 inhibition facilitates urinary neutral amino acid excretion and lowers plasma phenylalanine. JCI Insight. 2024 Nov 8;9(21):e182876.

140. Schulze RJ, Strom SC, Nyberg SL. From pain to gain: Leveraging acetaminophen in hepatocyte transplantation for phenylketonuria. Hepatology. 2024 May;79(5):973–5.

141. Martinez M, Harding CO, Schwank G, Thöny B. State-of-the-art 2023 on gene therapy for phenylketonuria. J of Inher Metab Disea. 2024 Jan;47(1):80–92.

142. Iansante V, Mitry RR, Filippi C, Fitzpatrick E, Dhawan A. Human hepatocyte transplantation for liver disease: current status and future perspectives. Pediatr Res. 2018 Jan;83(1–2):232–40.

143. Vonada A, Wakefield L, Martinez M, Harding CO, Grompe M, Tiyaboonchai A. Complete correction of murine phenylketonuria by selection-enhanced hepatocyte transplantation. Hepatology. 2024 May;79(5):1088–97.

144. Baek R, Coughlan K, Jiang L, Liang M, Ci L, Singh H, et al. Characterizing the mechanism of action for mRNA therapeutics for the treatment of propionic acidemia, methylmalonic acidemia, and phenylketonuria. Nat Commun. 2024 May 7;15(1):3804.

145. Perez-Garcia CG, Diaz-Trelles R, Vega JB, Bao Y, Sablad M, Limphong P, et al. Development of an mRNA replacement therapy for phenylketonuria. Molecular Therapy - Nucleic Acids. 2022 Jun;28:87–98.

146. Cacicedo ML, Weinl-Tenbruck C, Frank D, Limeres MJ, Wirsching S, Hilbert K, et al. Phenylalanine hydroxylase mRNA rescues the phenylketonuria phenotype in mice. Front Bioeng Biotechnol. 2022 Oct 7;10:993298.

147. Diaz-Trelles R, Lee S, Kuakini K, Park J, Dukanovic A, Gonzalez JA, et al. Lipid nanoparticle delivers phenylalanine ammonia lyase mRNA to the liver leading to catabolism and clearance of phenylalanine in a phenylketonuria mouse model. Molecular Genetics and Metabolism Reports. 2022 Sep;32:100882.

148. O'Leary E, Jiang Y, Kristensen LS, Hansen TB, Kjems J. The therapeutic potential of circular RNAs. Nat Rev Genet. 2025 Apr;26(4):230–44.

149. Rebuffat A, Harding CO, Ding Z, Thöny B. Comparison of Adeno-Associated Virus Pseudotype 1, 2, and 8 Vectors Administered by Intramuscular Injection in the Treatment of Murine Phenylketonuria. Human Gene Therapy. 2010 Apr;21(4):463–77.

150. Ahmed SS, Rubin H, Wang M, Faulkner D, Sengooba A, Dollive SN, et al. Sustained Correction of a Murine Model of Phenylketonuria following a Single Intravenous Administration of AAVHSC15-PAH. Molecular Therapy - Methods & Clinical Development. 2020 Jun;17:568–80.

151. Vonada A, Tiyaboonchai A, Nygaard S, Posey J, Peters AM, Winn SR, et al. Therapeutic liver repopulation by transient acetaminophen selection of gene-modified hepatocytes. Sci Transl Med. 2021 Jun 9;13(597):eabg3047.

152. B. Fang, R. C. Eisensmith, X. H. C. Li, M. J. Finegold, A. Shedlovsky, W. Dove, S. L. C. Woo. Gene therapy for phenylketonuria: phenotypic correction in a genetically deficient mouse model by adenovirus-mediated hepatic gene transfer. Gene Therapy. 1(4):247–54.

153. Oh HJ, Park ES, Kang S, Jo I, Jung SC. Long-Term Enzymatic and Phenotypic Correction in the Phenylketonuria Mouse Model by Adeno-Associated Virus Vector-Mediated Gene Transfer. Pediatr Res. 2004 Aug;56(2):278–84.

154. Mochizuki S, Mizukami H, Ogura T, Kure S, Ichinohe A, Kojima K, et al. Long-term correction of hyperphenylalaninemia by AAV-mediated gene transfer leads to behavioral recovery in phenylketonuria mice. Gene Ther. 2004 Jul 1;11(13):1081–6.

155. Ding Z, Georgiev P, Thöny B. Administration-route and gender-independent long-term therapeutic correction of phenylketonuria (PKU) in a mouse model by recombinant adeno-associated virus 8 pseudotyped vector-mediated gene transfer. Gene Ther. 2006 Apr 1;13(7):587–93.

156. Ding Z, Harding CO, Rebuffat A, Elzaouk L, Wolff JA, Thöny B. Correction of Murine PKU Following AAV-mediated Intramuscular Expression of a Complete Phenylalanine Hydroxylating System. Molecular Therapy. 2008 Apr;16(4):673–81.

157. Harding CO, Gillingham MB, Hamman K, Clark H, Goebel-Daghighi E, Bird A, et al. Complete correction of hyperphenylalaninemia following liver-directed, recombinant AAV2/8 vector-mediated gene therapy in murine phenylketonuria. Gene Ther. 2006 Mar 1;13(5):457–62.

158. BioSpace. Homology Medicines reports second quarter 2023 financial results and recent highlights.

159. BioMarin Pharmaceutical Inc. BioMarin announces fourth quarter and full year 2021 financial results and corporate updates. 2022 [Internet]. Available from: https://www.biomarin.com/news/press-releases/biomarin-announces-fourth-quarter-and-full-year-2021-financial-results-and-corporate-updates/

160. Nelson CE, Wu Y, Gemberling MP, Oliver ML, Waller MA, Bohning JD, et al. Long-term evaluation of AAV-CRISPR genome editing for Duchenne muscular dystrophy. Nat Med. 2019 Mar;25(3):427–32.

161. Husain T, Passini MA, Parente MK, Fraser NW, Wolfe JH. Long-term AAV vector gene and protein expression in mouse brain from a small pan-cellular promoter is similar to neural cell promoters. Gene Ther. 2009 Jul;16(7):927–32.

162. Coppoletta JM. BODY LENGTH AND ORGAN WEIGHTS OF INFANTS AND CHILDREN * A STUDY OF THE BODY LENGTH AND NORMAL WEIGHTS OF THE MORE IMPORTANT VITAL ORGANS OF THE BODY BE- TWEEN BIRTH AND TWELVE YEARS OF AGE.

163. Cunningham SC, Dane AP, Spinoulas A, Alexander IE. Gene Delivery to the Juvenile Mouse Liver Using AAV2/8 Vectors. Molecular Therapy. 2008 Jun;16(6):1081–8.

164. Heinke P, Rost F, Rode J, Trus P, Simonova I, Lázár E, et al. Diploid hepatocytes drive physiological liver renewal in adult humans. Cell Systems. 2022 Jun;13(6):499-507.e12.

165. Niemeyer GP, Herzog RW, Mount J, Arruda VR, Tillson DM, Hathcock J, et al. Long-term correction of inhibitor-prone hemophilia B dogs treated with liver-directed AAV2-mediated factor IX gene therapy. Blood. 2009 Jan 22;113(4):797–806.

166. Nathwani AC, Reiss UM, Tuddenham EGD, Rosales C, Chowdary P, McIntosh J, et al. Long-Term Safety and Efficacy of Factor IX Gene Therapy in Hemophilia B. N Engl J Med. 2014 Nov 20;371(21):1994–2004.

167. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology. 2005 Aug 1;151(8):2551–61.

168. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. Science. 2007 Mar 23;315(5819):1709–12.

169. Mojica FJM, D□ez-Villase□or C, Garc□a-Mart□nez J, Soria E. Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. J Mol Evol. 2005 Feb;60(2):174–82.

170. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science. 2012 Aug 17;337(6096):816–21.

171. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. Science. 2013 Feb 15;339(6121):819–23.

172. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-Guided Human Genome Engineering via Cas9. Science. 2013 Feb 15;339(6121):823–6.

173. Friedland AE, Baral R, Singhal P, Loveluck K, Shen S, Sanchez M, et al. Characterization of Staphylococcus aureus Cas9: a smaller Cas9 for all-in-one adeno-associated virus delivery and paired nickase applications. Genome Biol. 2015 Dec;16(1):257.

174. Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, et al. In vivo genome editing using Staphylococcus aureus Cas9. Nature. 2015 Apr;520(7546):186–91.

175. Fonfara I, Richter H, Bratovič M, Le Rhun A, Charpentier E. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. Nature. 2016 Apr;532(7600):517–21.

176. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, et al. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. Cell. 2013 Feb;152(5):1173–83.

177. Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, et al. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. Cell. 2013 Jul;154(2):442–51.

178. Suzuki K, Izpisua Belmonte JC. In vivo genome editing via the HITI method as a tool for gene therapy. J Hum Genet. 2018 Feb;63(2):157–64.

179. Lieber MR. The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway. Annu Rev Biochem. 2010 Jun 7;79(1):181–211.

180. Chen W, McKenna A, Schreiber J, Haeussler M, Yin Y, Agarwal V, et al. Massively parallel profiling and predictive modeling of the outcomes of CRISPR/Cas9-mediated double-strand break repair. Nucleic Acids Research. 2019 Sep 5;47(15):7989–8003.

181. Suzuki K, Tsunekawa Y, Hernandez-Benitez R, Wu J, Zhu J, Kim EJ, et al. In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. Nature. 2016 Dec 1;540(7631):144–9.

182. Uddin F, Rudin CM, Sen T. CRISPR Gene Therapy: Applications, Limitations, and Implications for the Future. Front Oncol. 2020 Aug 7;10:1387.

183. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature. 2016 May 19;533(7603):420–4.

184. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. Nature. 2017 Nov 23;551(7681):464–71.

185. Kim YB, Komor AC, Levy JM, Packer MS, Zhao KT, Liu DR. Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. Nat Biotechnol. 2017 Apr;35(4):371–6.

186. Anzalone AV, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, et al. Search-and-replace genome editing without double-strand breaks or donor DNA. Nature. 2019 Dec 5;576(7785):149–57.

187. Chen A, Pan Y, Chen J. Clinical, genetic, and experimental research of hyperphenylalaninemia. Front Genet. 2023 Jan 4;13:1051153.

188. Anzalone AV, Koblan LW, Liu DR. Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. Nat Biotechnol. 2020 Jul 1;38(7):824–44.

189. Schneller JL, Lee CM, Venturoni LE, Chandler RJ, Li A, Myung S, et al. In vivo genome editing at the albumin locus to treat methylmalonic acidemia. Molecular Therapy - Methods & Clinical Development. 2021 Dec;23:619–32.

190. Sharma R, Anguela XM, Doyon Y, Wechsler T, DeKelver RC, Sproul S, et al. In vivo genome editing of the albumin locus as a platform for protein replacement therapy. Blood. 2015 Oct 8;126(15):1777–84.

191. Chen HM, Resendes R, Ghodssi A, Sookiasian D, Tian M, Dollive S, et al. Molecular characterization of precise in vivo targeted gene integration in human cells using AAVHSC15. Lewin AS, editor. PLoS ONE. 2020 May 26;15(5):e0233373.

192. Villiger L, Grisch-Chan HM, Lindsay H, Ringnalda F, Pogliano CB, Allegri G, et al. Treatment of a metabolic liver disease by in vivo genome base editing in adult mice. Nat Med. 2018 Oct;24(10):1519–25.

193. Villiger L, Rothgangl T, Witzigmann D, Oka R, Lin PJC, Qi W, et al. In vivo cytidine base editing of hepatocytes without detectable off-target mutations in RNA and DNA. Nat Biomed Eng. 2021 Jan 25;5(2):179–89.

194. Brooks DL, Whittaker MN, Said H, Dwivedi G, Qu P, Musunuru K, et al. A base editing strategy using mRNA-LNPs for in vivo correction of the most frequent phenylketonuria variant. Human Genetics and Genomics Advances. 2024 Jan;5(1):100253.

195. Böck D, Rothgangl T, Villiger L, Schmidheini L, Matsushita M, Mathis N, et al. In vivo prime editing of a metabolic liver disease in mice. Sci Transl Med. 2022 Mar 16;14(636):eabl9238.

196. Rothgangl T, Tálas A, Ioannidi EI, Weber Y, Böck D, Matsushita M, et al. Treatment of a metabolic liver disease in mice with a transient prime editing approach. Nat Biomed Eng [Internet]. 2025 May 20 [cited 2025 Aug 31]; Available from: https://www.nature.com/articles/s41551-025-01399-4

197. Tessera Therapeutics. Tessera Therapeutics features new preclinical data demonstrating progress across its in vivo Gene Writing™ programs and delivery platform at the American Society of Gene and Cell Therapy 28th annual meeting. 2025 [Internet]. Available from: https://www.tesseratherapeutics.com/news/tessera-therapeutics-features-new-preclinical-data-demonstrating-progress-across-its-in-vivo-gene-writing-programs-and-delivery-platform-at-the-american-society-of-gene-and-cell-therapy-28th-annual-meeting?utm_source=chatgpt.com

198. Kaufman S. The Phenylalanine Hydroxylating System in Phenylketonuria and Its Variants.

199. Berry HK. Diagnosis of Phenylalanine Hydroxylase Deficiency (Phenylketonuria). Arch Pediatr Adolesc Med. 1982 Feb 1;136(2):111.

200. Grimm U, Knapp A, Schlenzka K, Hesse R. Phenylalanine hydroxylase activity in heterozygote hereditary carriers of the phenylketonuria gene. Acta Biol Med Ger. 1977;36((7-8):1179–82.

201. Hamman K, Clark H, Montini E, Al-Dhalimy M, Grompe M, Finegold M, et al. Low Therapeutic Threshold for Hepatocyte Replacement in Murine Phenylketonuria. Molecular Therapy. 2005 Aug;12(2):337–44.

202. Hamman KJ, Winn SR, Harding CO. Hepatocytes from wild-type or heterozygous donors are equally effective in achieving successful therapeutic liver repopulation in murine phenylketonuria (PKU). Molecular Genetics and Metabolism. 2011 Nov;104(3):235–40.

203. Elzaouk L, Laufs S, Heerklotz D, Leimbacher W, Blau N, Résibois A, et al. Nuclear localization of tetrahydrobiopterin biosynthetic enzymes. Biochimica et Biophysica Acta (BBA) - General Subjects. 2004 Jan 5;1670(1):56–68.

204. Truett GE, Heeger P, Mynatt RL, Truett AA, Walker JA, Warman ML. Preparation of PCR-Quality Mouse Genomic DNA with Hot Sodium Hydroxide and Tris (HotSHOT). BioTechniques. 2000 Jul;29(1):52–4.

205. Graham FL, Smiley J, Nairn R. Characteristics of a Human Cell Line Transformed by DNA from Human Adenovirus Type 5. Journal of General Virology. 1977 Jul 1;36(1):59–72.

206. Todaro GJ, Green H. QUANTITATIVE STUDIES OF THE GROWTH OF MOUSE EMBRYO CELLS IN CULTURE AND THEIR DEVELOPMENT INTO ESTABLISHED LINES. The Journal of Cell Biology. 1963 May 1;17(2):299–313.

207. Otey CA, Boukhelifa M, Maness P. B35 Neuroblastoma Cells: An Easily Transfected, Cultured Cell Model of Central Nervous System Neurons. In: Methods in Cell Biology [Internet]. Elsevier;

2003 [cited 2025 Aug 31]. p. 287–304. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0091679X03010136

208. Gao GP, Alvira MR, Wang L, Calcedo R, Johnston J, Wilson JM. Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. Proc Natl Acad Sci USA. 2002 Sep 3;99(18):11854–9.

209. Miciak JJ, Hirshberg J, Bunz F. Seamless assembly of recombinant adenoviral genomes from high-copy plasmids. Ulasov I, editor. PLoS ONE. 2018 Jun 27;13(6):e0199563.

210. Ginn S, Doroudian D. Highly Efficient Functional Editing of the Native Ornithine Transcarbamylase Locus by Targeted Integration with Phenotype Correction and Restoration of Physiological Patterns of Gene Expression,. PREPRINT (Version 1) available at Research Square.

211. Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, et al. In vivo genome editing using Staphylococcus aureus Cas9. Nature. 2015 Apr;520(7546):186–91.

212. Gustafsson C, Govindarajan S, Minshull J. Codon bias and heterologous protein expression. Trends in Biotechnology. 2004 Jul;22(7):346–53.

213. Ginn S, Doroudian F, Christina S, Chan O, Lucas C, Zhu E, et al. Highly Efficient Functional Editing of the Native Ornithine Transcarbamylase Locus by Targeted Integration with Phenotype Correction and Restoration of Physiological Patterns of Gene Expression [Internet]. In Review; 2025 [cited 2026 Jan 29]. Available from: https://www.researchsquare.com/article/rs-7237670/v1

214. Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, et al. QuPath: Open source software for digital pathology image analysis. Sci Rep. 2017 Dec 4;7(1):16878.

215. Wang L, Wang H, Bell P, McCarter RJ, He J, Calcedo R, et al. Systematic Evaluation of AAV Vectors for Liver directed Gene Transfer in Murine Models. Molecular Therapy. 2010 Jan;18(1):118–25.

216. Gao G, Lu Y, Calcedo R, Grant RL, Bell P, Wang L, et al. Biology of AAV Serotype Vectors in Liver-Directed Gene Transfer to Nonhuman Primates. Molecular Therapy. 2006 Jan;13(1):77–87.

217. Ebadat S, Ahmadi S, Ahmadi M, Nematpour F, Barkhordari F, Mahdian R, et al. Evaluating the efficiency of CHEF and CMV promoter with IRES and Furin/2A linker sequences for monoclonal antibody expression in CHO cells. Stieger K, editor. PLoS ONE. 2017 Oct 12;12(10):e0185967.

218. Mizuguchi H, Xu Z, Ishii-Watabe A, Uchida E, Hayakawa T. IRES-Dependent Second Gene Expression Is Significantly Lower Than Cap-Dependent First Gene Expression in a Bicistronic Vector. Molecular Therapy. 2000 Apr;1(4):376–82.

219. Bouabe H, Fassler R, Heesemann J. Improvement of reporter activity by IRES-mediated polycistronic reporter system. Nucleic Acids Research. 2008 Jan 10;36(5):e28–e28.

220. Donnelly MLL, Hughes LE, Luke G, Mendoza H, Ten Dam E, Gani D, et al. The ‘cleavage’ activities of foot-and-mouth disease virus 2A site-directed mutants and naturally occurring ‘2A-like’ sequences. Journal of General Virology. 2001 May 1;82(5):1027–41.

221. Liu Z, Chen O, Wall JBJ, Zheng M, Zhou Y, Wang L, et al. Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. Sci Rep. 2017 May 19;7(1):2193.

222. Kim JH, Lee SR, Li LH, Park HJ, Park JH, Lee KY, et al. High Cleavage Efficiency of a 2A Peptide Derived from Porcine Teschovirus-1 in Human Cell Lines, Zebrafish and Mice. Thiel V, editor. PLoS ONE. 2011 Apr 29;6(4):e18556.

223. Harding CO, Gillingham MB, Hamman K, Clark H, Goebel-Daghighi E, Bird A, et al. Complete correction of hyperphenylalaninemia following liver-directed, recombinant AAV2/8 vector-mediated gene therapy in murine phenylketonuria. Gene Ther. 2006 Mar 1;13(5):457–62.

224. Concordet JP, Haeussler M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. Nucleic Acids Research. 2018 Jul 2;46(W1):W242–5.

225. Tycko J, Barrera LA, Huston NC, Friedland AE, Wu X, Gootenberg JS, et al. Pairwise library screen systematically interrogates Staphylococcus aureus Cas9 specificity in human cells. Nat Commun. 2018 Jul 27;9(1):2962.

226. Najm FJ, Strand C, Donovan KF, Hegde M, Sanson KR, Vaimberg EW, et al. Orthologous CRISPR–Cas9 enzymes for combinatorial genetic screens. Nat Biotechnol. 2018 Feb;36(2):179–89.

227. Synthego. 2023.

228. Ceccaldi R, Rondinelli B, D'Andrea AD. Repair Pathway Choices and Consequences at the Double-Strand Break. Trends in Cell Biology. 2016 Jan;26(1):52–64.

229. Bennett EP, Petersen BL, Johansen IE, Niu Y, Yang Z, Chamberlain CA, et al. INDEL detection, the 'Achilles heel' of precise genome editing: a survey of methods for accurate profiling of gene editing induced indels. Nucleic Acids Research. 2020 Dec 2;48(21):11958–81.

230. Gustafsson C, Govindarajan S, Minshull J. Codon bias and heterologous protein expression. Trends in Biotechnology. 2004 Jul;22(7):346–53.

231. Moorman AFM, Vermeulen JLM, Charles R, Lamers WH. Localization of ammonia-metabolizing enzymes in human liver: Ontogenesis of heterogeneity. Hepatology. 1989 Mar;9(3):367–72.

232. Heintz C, Troxler H, Martinez A, Thöny B, Blau N. Quantification of phenylalanine hydroxylase activity by isotope-dilution liquid chromatography–electrospray ionization tandem mass spectrometry. Molecular Genetics and Metabolism. 2012 Apr;105(4):559–65.

233. Luqman MW, Jenjaroenpun P, Spathos J, Shingte N, Cummins M, Nimsamer P, et al. Long read sequencing reveals transgene concatemerization and vector sequences integration following AAV-driven electroporation of CRISPR RNP complexes in mouse zygotes. Front Genome Ed. 2025 Jun 4;7:1582097.

234. Suchy FP, Karigane D, Nakauchi Y, Higuchi M, Zhang J, Pekrun K, et al. Genome engineering with Cas9 and AAV repair templates generates frequent concatemeric insertions of viral vectors. Nat Biotechnol. 2025 Feb;43(2):204–13.

235. Wu Q, Bazzini AA. Translation and mRNA Stability Control. Annu Rev Biochem. 2023 Jun 20;92(1):227–45.

236. Halpern KB, Shenhav R, Matcovitch-Natan O, Tóth B, Lemze D, Golan M, et al. Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. Nature. 2017 Feb 16;542(7641):352–6.

237. Ying M, Pey AL, Aarsæther N, Martinez A. Phenylalanine hydroxylase expression in primary rat hepatocytes is modulated by oxygen concentration. Molecular Genetics and Metabolism. 2010 Oct;101(2–3):279–81.

238. Aizarani N, Saviano A, Sagar, Mailly L, Durand S, Herman JS, et al. A human liver cell atlas reveals heterogeneity and epithelial progenitors. Nature. 2019 Aug;572(7768):199–204.

239. Yu S, Wang H, Yang L, Yan Y, Cai Q, Ma D, et al. Spatial transcriptome profiling of normal human liver. Sci Data. 2022 Oct 19;9(1):633.

240. Watson BR, Paul B, Rahman RU, Amir-Zilberstein L, Segerstolpe Å, Epstein ET, et al. Spatial transcriptomics of healthy and fibrotic human liver at single-cell resolution. Nat Commun. 2025 Jan 2;16(1):319.

241. Ben-Moshe S. Spatial sorting enables comprehensive characterization of liver zonation. 2019;

242. Cunningham SC, Spinoulas A, Carpenter KH, Wilcken B, Kuchel PW, Alexander IE. AAV2/8-mediated Correction of OTC Deficiency Is Robust in Adult but Not Neonatal Spfash Mice. Molecular Therapy. 2009 Aug;17(8):1340–6.

243. Dane AP, Cunningham SC, Graf NS, Alexander IE. Sexually Dimorphic Patterns of Episomal rAAV Genome Persistence in the Adult Mouse Liver and Correlation With Hepatocellular Proliferation. Molecular Therapy. 2009 Sep;17(9):1548–54.

244. Bell P, Wang L, Gao G, Haskins ME, Tarantal AF, McCarter RJ, et al. Inverse zonation of hepatocyte transduction with AAV vectors between mice and non-human primates. Molecular Genetics and Metabolism. 2011 Nov;104(3):395–403.

245. Cabanes-Creus M, Navarro RG, Liao SHY, Scott S, Carlessi R, Roca-Pinilla R, et al. Characterization of the humanized FRG mouse model and development of an AAV-LK03 variant with improved liver lobular biodistribution. Molecular Therapy - Methods & Clinical Development. 2023 Mar;28:220–37.

246. Lisowski L, Dane AP, Chu K, Zhang Y, Cunningham SC, Wilson EM, et al. Selection and evaluation of clinically relevant AAV variants in a xenograft liver model. Nature. 2014 Feb;506(7488):382–6.

247. Zhang MH, Liu XM, Zhang C. Random Integration Analysis of Recombinant Adeno-Associated Virus 6 Packaged in Sf9 Insect Cells. Mol Biol. 2023 Aug;57(4):724–34.

248. Haberman RP, McCown TJ, Samulski RJ. Novel Transcriptional Regulatory Signals in the Adeno-Associated Virus Terminal Repeat *A/D* Junction Element. J Virol. 2000 Sep 15;74(18):8732–9.

249. Earley LF, Conatser LM, Lue VM, Dobbins AL, Li C, Hirsch ML, et al. Adeno-Associated Virus Serotype-Specific Inverted Terminal Repeat Sequence Role in Vector Transgene Expression. Human Gene Therapy. 2020 Feb 1;31(3–4):151–62.

250. Choi TB, Pardridge WM. Phenylalanine transport at the human blood-brain barrier. Studies with isolated human brain capillaries. Journal of Biological Chemistry. 1986 May;261(14):6536–41.

251. Luo L, Zeng H, Hu Y, Jiang L, Fu C, Huang J, et al. The amino acid transporter SLC16A10 promotes melanogenesis by facilitating the transportation of phenylalanine. Experimental Dermatology. 2024 Aug;33(8):e15165.

252. Ertl HCJ. Immunogenicity and toxicity of AAV gene therapy. Front Immunol. 2022 Aug 12;13:975803.

253. Huang, H-R. ITherapeutics. CRISPR/Cas9-Mediated Targeted Insertion of Human F9 Achieves Therapeutic Circulating Protein Levels in Mice and Non-Human Primates. Oral presentation presented at: 22nd Annual Meeting of the American Society of Gene & Cell Therapy (ASGCT),; 2019 Apr 29.

254. Li S, Hu Y, Li A, Lin J, Hsieh K, Schneiderman Z, et al. Payload distribution and capacity of mRNA lipid nanoparticles. Nat Commun. 2022 Sep 23;13(1):5561.

255. Han JP, Kim M, Choi BS, Lee JH, Lee GS, Jeong M, et al. In vivo delivery of CRISPR-Cas9 using lipid nanoparticles enables antithrombin gene editing for sustainable hemophilia A and B therapy. Sci Adv. 2022 Jan 21;8(3):eabj6901.

256. Finn JD, Smith AR, Patel MC, Shaw L, Youniss MR, Van Heteren J, et al. A Single Administration of CRISPR/Cas9 Lipid Nanoparticles Achieves Robust and Persistent In Vivo Genome Editing. Cell Reports. 2018 Feb;22(9):2227–35.

257. Chen K, Han H, Zhao S, Xu B, Yin B, Lawanprasert A, et al. Lung and liver editing by lipid nanoparticle delivery of a stable CRISPR–Cas9 ribonucleoprotein. Nat Biotechnol [Internet]. 2024 Oct 16 [cited 2025 Aug 31]; Available from: https://www.nature.com/articles/s41587-024-02437-3

258. Kenjo E, Hozumi H, Makita Y, Iwabuchi KA, Fujimoto N, Matsumoto S, et al. Low immunogenicity of LNP allows repeated administrations of CRISPR-Cas9 mRNA into skeletal muscle in mice. Nat Commun. 2021 Dec 8;12(1):7101.

259. Yin H, Song CQ, Suresh S, Wu Q, Walsh S, Rhym LH, et al. Structure-guided chemical modification of guide RNA enables potent non-viral in vivo genome editing. Nat Biotechnol. 2017 Dec;35(12):1179–87.

260. Locatelli F, Lang P, Wall D, Meisel R, Corbacioglu S, Li AM, et al. Exagamglogene Autotemcel for Transfusion-Dependent β-Thalassemia. N Engl J Med. 2024 May 9;390(18):1663–76.

261. Rosenblum D, Gutkin A, Kedmi R, Ramishetti S, Veiga N, Jacobi AM, et al. CRISPR-Cas9 genome editing using targeted lipid nanoparticles for cancer therapy. Sci Adv. 2020 Nov 20;6(47):eabc9450.

262. Yin H, Song CQ, Dorkin JR, Zhu LJ, Li Y, Wu Q, et al. Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo. Nat Biotechnol. 2016 Mar;34(3):328–33.

263. Foley RA, Ayoub PG, Sinha V, Juett C, Sanoyca A, Duggan EC, et al. Lipid Nanoparticles for the Delivery of CRISPR/Cas9 Machinery to Enable Site-Specific Integration of CFTR and Mutation-Agnostic Disease Rescue [Internet]. Bioengineering; 2025 [cited 2025 Aug 31]. Available from: http://biorxiv.org/lookup/doi/10.1101/2025.01.22.633938

264. Musunuru K, Grandinette SA, Wang X, Hudson TR, Briseno K, Berry AM, et al. Patient-Specific In Vivo Gene Editing to Treat a Rare Genetic Disease. N Engl J Med. 2025 Jun 12;392(22):2235–43.

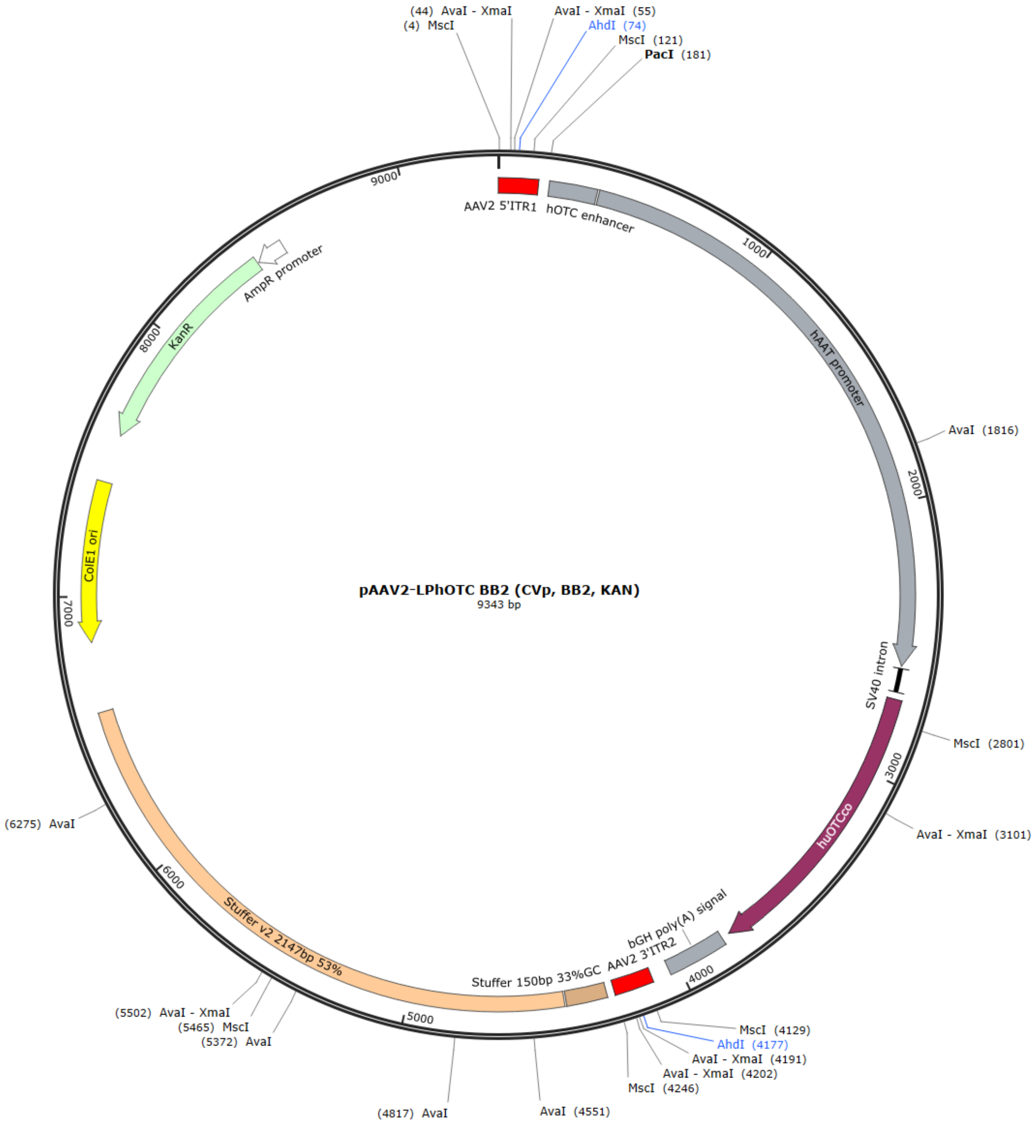
`

6. Appendices

6.1 Methods and Analysis tools

6.1.1 pAAV2-LPhOTC BB2 backbone plasmid map

Self-complimentary rAAV2 backbone plasmid from Assoc/Prof Samantha Ginn.



6.1.2 Synthesised donor DNA sequences

Appendices Table 6 1: Synthesised donor DNA sequences.

Donor name	Sequence 5’-3’
Hybrid GFP	TTAATTAAATTCTTGGTGCATATCGGCATGTGCTACATAGTACATGGGTCTTTTCTGACATACATATTTCTCCCTTTTAAATCCTTCTTTTTCAGATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGATATCGTCGACTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACATTCTTGGTGCATATCGGCATGTGCTAACTAGT
Universal GFP	CTTAATTAAATTCTTGGTGCATATCGGCATGTGCTACAATGAATCACCATAGTACATGGGTCTTTTCTGAAATACATATTTCTCCCTTTTAAATCTCTTTTTACAGGGAAGCGGAGCTACTAACTTCAGCCTGCTGAAGCAGGCTGGAGACGTGGAGGAGAACCCTATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGATATCGTCGACTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACATTCTTGGTGCATATCGGCATGTGCTAACTAGT
Hybrid hcoPAH	TTAATTAAATTCTTGGTGCATATCGGCATGTGCTAcatagtacatgggtcttttctgacatacatatttctcccttttaaatccttctttttcagGAGACCTCCTACATCGAGGACAATTGTAACCAGAATGGCGCCATCAGTCTGATCTTCAGCTTAAAGGAGGAGGTGGGTGCCCTCGCCAAGGTTCTGCGGCTGTTTGAAGAAAATGACGTGAACCTGACCCACATTGAAAGCCGGCCTTCTCGGCTGAAGAAGGATGAGTATGAATTCTTCACACACCTGGACAAAAGATCCTTGCCAGCCCTGACTAACATCATCAAGATCTTGAGACATGACATCGGGGCCACAGTCCACGAGCTCTCCAGAGACAAGAAAAAGGACACAGTACCCTGGTTCCCTAGAACCATACAAGAACTGGACAGATTTGCCAACCAAATCCTGTCCTATGGCGCAGAGCTGGATGCTGACCACCCCGGATTCAAAGACCCTGTCTACCGAGCCCGAAGAAAGCAATTTGCTGATATCGCCTACAATTACAGGCACGGCCAGCCCATCCCTCGCGTGGAATACATGGAGGAAGAGAAGAAGACCTGGGGCACGGTATTCAAGACACTGAAAAGCCTCTACAAAACTCACGCCTGCTATGAGTACAACCACATCTTCCCTCTGCTGGAAAAATACTGTGGCTTCCATGAAGACAATATACCACAGCTGGAAGATGTGAGCCAGTTCCTCCAAACCTGCACAGGCTTCAGGCTCCGGCCTGTGGCTGGCCTGCTGAGCTCCCGGGACTTCCTGGGAGGCCTGGCCTTCCGGGTCTTCCACTGCACCCAGTATATCCGCCATGGCAGCAAGCCCATGTACACCCCAGAGCCTGACATCTGCCATGAGCTTTTGGGCCACGTCCCTCTTTTCAGTGACAGAAGCTTCGCTCAGTTTTCCCAGGAGATTGGCCTGGCCAGCCTGGGAGCCCCAGATGAATACATTGAGAAGCTAGCCACCATCTACTGGTTTACCGTGGAGTTTGGGCTGTGTAAACAAGGAGACTCCATAAAAGCCTACGGAGCCGGTTTACTGAGCTCGTTTGGAGAGCTGCAGTACTGCCTGTCTGAGAAACCCAAGCTGCTGCCCCTTGAGCTGGAGAAGACAGCTATACAGAACTACACAGTGACGGAGTTCCAGCCCCTGTACTATGTGGCTGAAAGTTTCAATGATGCCAAGGAAAAGGTGAGGAACTTTGCAGCTACGATCCCTAGACCCTTCTCTGTGAGATATGACCCCTACACTCAGCGCATTGAAGTGTTAGATAACACACAGCAGCTTAAAATCCTCGCAGATTCCATCAACTCTGAGATTGGCATTCTTTGTAGCGCTCTGCAGAAGATCAAGTGAGATatcgtcgactcgctgatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctcccccgtgccttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattctggggggtggggtggggcaggacagcaaggggga ggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggcttctgaggcggaaagaacATTCTTGGTGCATATCGGCATGTGCTAACTAGT

Universal hcoPAH	CTTAATTAAATTCTTGGTGCATATCGGCATGTGCTACAATGAATCACCATAGTACATGGGTCTTTTCTGAAATACATATTTCTCCCTTTTAAATCTCTTTTTACAGGGAAGCGGAGCTACTAACTTCAGCCTGCTGAAGCAGGCTGGAGACGTGGAGGAGAACCCTATGTCTACCGCTGTGCTGGAGAACCCTGGTCTGGGCAGAAAGCTCTCAGACTTTGGCCAGGAGACCTCCTACATCGAGGACAATTGTAACCAGAATGGCGCCATCAGTCTGATCTTCAGCTTAAAGGAGGAGGTGGGTGCCCTCGCCAAGGTTCTGCGGCTGTTTGAAGAAAATGACGTGAACCTGACCCACATTGAAAGCCGGCCTTCTCGGCTGAAGAAGGATGAGTATGAATTCTTCACACACCTGGACAAAAGATCCTTGCCAGCCCTGACTAACATCATCAAGATCTTGAGACATGACATCGGGGCCACAGTCCACGAGCTCTCCAGAGACAAGAAAAAGGACACAGTACCCTGGTTCCCTAGAACCATACAAGAACTGGACAGATTTGCCAACCAAATCCTGTCCTATGGCGCAGAGCTGGATGCTGACCACCCCGGATTCAAAGACCCTGTCTACCGAGCCCGAAGAAAGCAATTTGCTGATATCGCCTACAATTACAGGCACGGCCAGCCCATCCCTCGCGTGGAATACATGGAGGAAGAGAAGAAGACCTGGGGCACGGTATTCAAGACACTGAAAAGCCTCTACAAAACTCACGCCTGCTATGAGTACAACCACATCTTCCCTCTGCTGGAAAAATACTGTGGCTTCCATGAAGACAATATACCACAGCTGGAAGATGTGAGCCAGTTCCTCCAAACCTGCACAGGCTTCAGGCTCCGGCCTGTGGCTGGCCTGCTGAGCTCCCGGGACTTCCTGGGAGGCCTGGCCTTCCGGGTCTTCCACTGCACCCAGTATATCCGCCATGGCAGCAAGCCCATGTACACCCCAGAGCCTGACATCTGCCATGAGCTTTTGGGCCACGTCCCTCTTTTCAGTGACAGAAGCTTCGCTCAGTTTTCCCAGGAGATTGGCCTGGCCAGCCTGGGAGCCCCAGATGAATACATTGAGAAGCTAGCCACCATCTACTGGTTTACCGTGGAGTTTGGGCTGTGTAAACAAGGAGACTCCATAAAAGCCTACGGAGCCGGTTTACTGAGCTCGTTTGGAGAGCTGCAGTACTGCCTGTCTGAGAAACCCAAGCTGCTGCCCCTTGAGCTGGAGAAGACAGCTATACAGAACTACACAGTGACGGAGTTCCAGCCCCTGTACTATGTGGCTGAAAGTTTCAATGATGCCAAGGAAAAGGTGAGGAACTTTGCAGCTACGATCCCTAGACCCTTCTCTGTGAGATATGACCCCTACACTCAGCGCATTGAAGTGTTAGATAACACACAGCAGCTTAAAATCCTCGCAGATTCCATCAACTCTGAGATTGGCATTCTTTGTAGCGCTCTGCAGAAGATCAAGTGAGATATCGTCGACTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACATTCTTGGTGCATATCGGCATGTGCTAACTAGT

6.1.3 QuPath image analysis software GFP-positive cell counting macro

QuPath compatible macro. Saved as ".groovy" file type.

Purpose: Detects DAPI-stained nuclei, classifies GFP+ cells (AF488 channel) as positive/negative, and measures their distances relative to the annotated tissue boundary.

Context: Used in Section 2.9.7; supports cell quantification in Section 3.6.

Environment: QuPath 0.4.4 (Windows 11); Fluorescence mode; pixel classifier "eva3" for tissue annotation.

Inputs/Outputs:

- **Input:** Whole-slide fluorescence image with DAPI (nuclear) and AF488 (GFP) channels.
- **Output:** Cell detection objects classified as positive/negative; exported measurements include cell counts and distances to tissue annotations.

Key parameters:

`

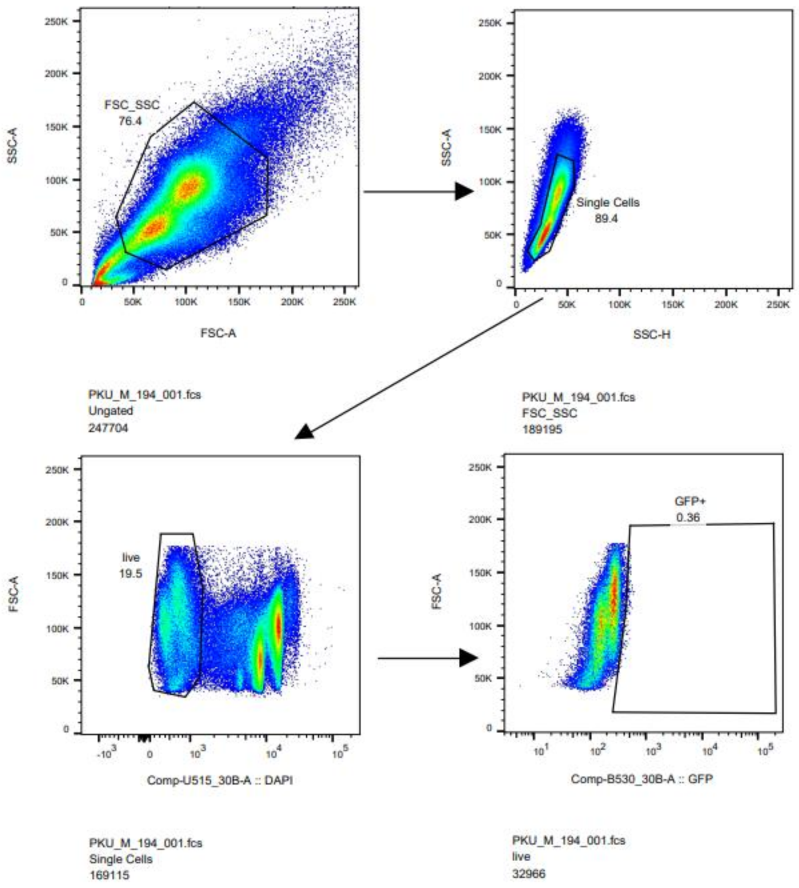
- Pixel size = 0.5 µm
- Nucleus area = 10–200 µm²
- Cytoplasmic AF488 threshold = 2000.0 (single threshold mode)
- Cell expansion = 5 µm

Provenance: cell_detection_distances.groovy (v1.0, 2025-08-27).

```
// outline the section
resetSelection();
setImageType('FLUORESCENCE');
createAnnotationsFromPixelClassifier("eva3", 0.0, 0.0)

// Select the "tissue" annotation
selectObjectsByClassification("tissue");

// detect positive and negative cells
runPlugin('qupath.imagej.detect.cells.PositiveCellDetection',
'{"detectionImage":"DAPI","requestedPixelSizeMicrons":0.5,"backgroundRadiusMicrons":8.0,"backgroundByReconstruction":true,"medianRadiusMicrons":0.5,"sigmaMicrons":0.6,"minAreaMicrons":10.0,"maxAreaMicrons":200.0,"threshold":1100.0,"watershedPostProcess":true,"cellExpansionMicrons":5.0,"includeNuclei":true,"smoothBoundaries":true,"makeMeasurements":true,"thresholdCompartment":"Cytoplasm: AF488 mean","thresholdPositive1":2000.0,"thresholdPositive2":2000.0,"thresholdPositive3":2000.0,"singleThreshold":true}');

// measure their distances from the annotations
detectionToAnnotationDistancesSigned(false)
resetSelection();
```

6.1.4 Flow cytometry gating strategy

Gating strategy employed for flow cytometry. Used in Section 2.9.5, supports flow analysis in Section 3.6.



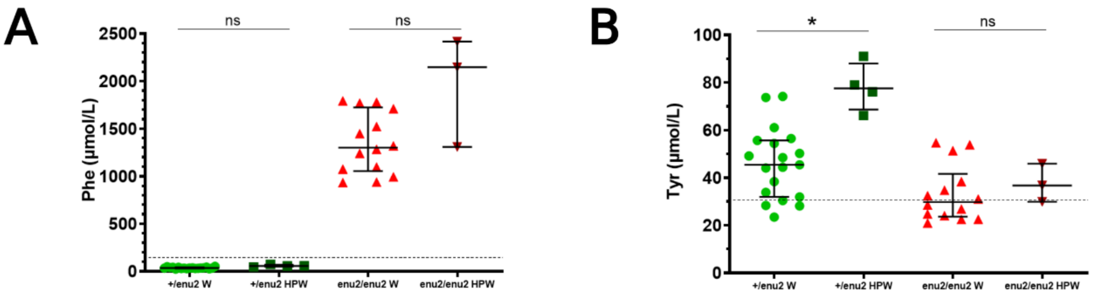
Appendices Figure 6.1: Flow cytometry gating strategy for GFP+ hepatocyte detection.

Representative gating workflow showing sequential identification of GFP-positive hepatocytes. Forward scatter (FSC-A) versus side scatter (SSC-A) was used to exclude debris and select the main cell population (top left). Singlets were gated based on FSC-A versus SSC-H to exclude doublets (top right). Live cells were identified by exclusion of DAPI-positive events (bottom left). Finally, GFP+ cells were quantified based on AF488 signal intensity (bottom right). Percentages of events retained at each gate are indicated. Flow cytometry and data analysis performed by Cindy Zhu.

6.2 Supporting Data

6.2.1 Phenylalanine supplementary drinking water

Additional way to increase dietary phenylalanine in Pah$^{enu2/enu2}$ mice. Relevant to Pahenu2 model phenotype characterisation (Section 3.2.1, Figure 3.1).



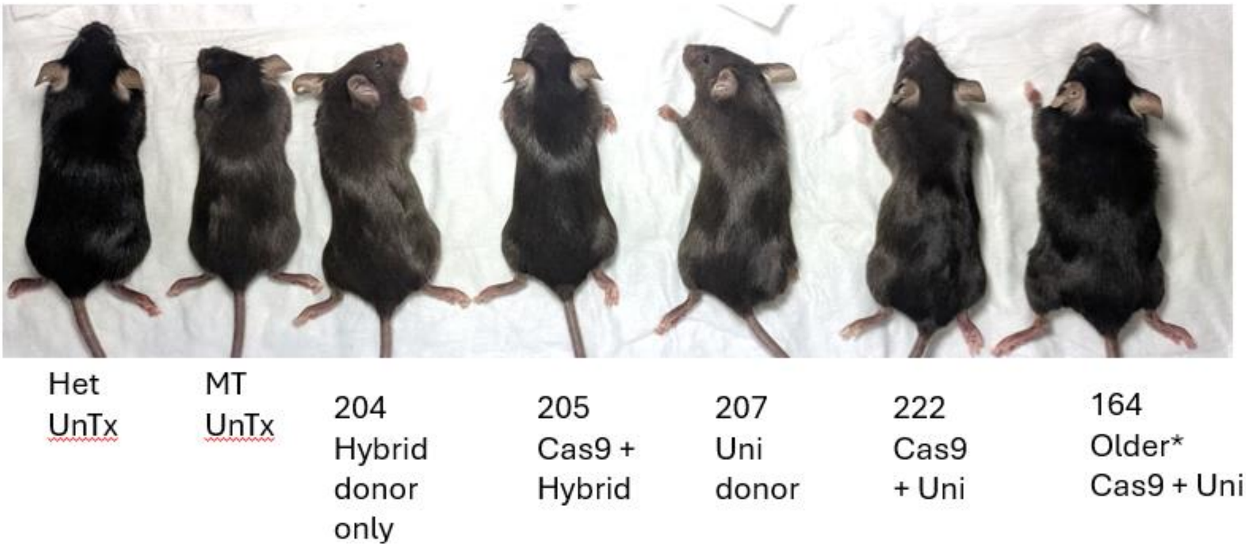
Appendices Figure 6.2: Blood amino acid levels in heterozygous and homozygous Pahenu2 mice when water supplemented with phenylalanine.

(A) Plasma phenylalanine (Phe) concentrations and (B) plasma tyrosine (Tyr) concentrations measured in heterozygous and homozygous *Pah*^{enu2} mice. “W” cohort on plain water while “HPW” on high phenylalanine supplemented water. Phenylalanine is not very soluble in water, but water is approximately 0.05% Phenylalanine (~50 mg/100mL). Each data point represents an individual animal; horizontal bars indicate median and interquartile range. Statistical comparisons were performed using one-way Kruskal Wallis with post hoc testing; $p < 0.05$ is indicated by *, ns = not significant. Dashed horizontal lines in (A) represent upper limit of healthy range while in (B) represent lower healthy limit.

`

6.2.2 Additional photos of coat colour over time

Extra images for Section 2.11 and Figure 3.10.

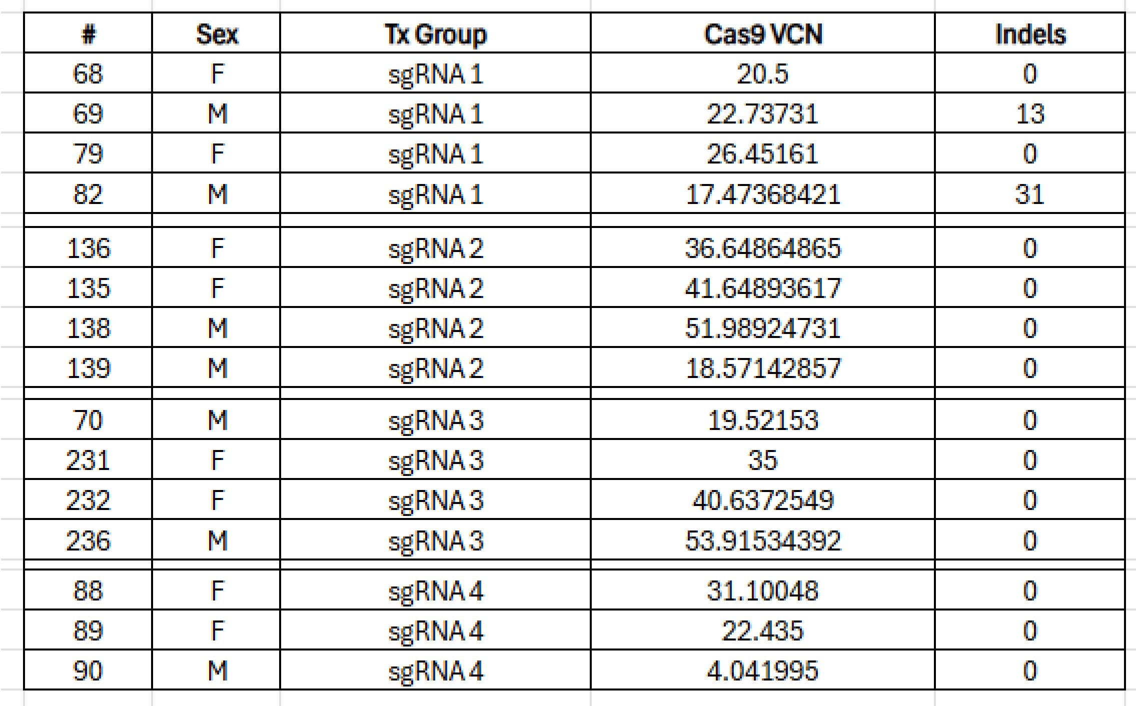


Appendices Figure 6.3: Representative additional images of fur phenotype reversion following dual CRISPR/Cas9 and PAH donor vector treatment.

Mice shown include untreated heterozygous (Het UnTx) and homozygous mutant (MT UnTx) controls, donor-only (Hybrid donor only; Uni donor only), and treated animals receiving Cas9 with hybrid or uni donor constructs. All mice, except for #164, are 16.5 weeks of age, approximately 8.5 weeks post injection. An older treated mouse (164; Cas9 + Uni), who is 37 weeks of age and 29 weeks post injection is also shown. Images illustrate partial restoration of normal black fur pigmentation compared to untreated mutant control.

`

6.2.3 All raw matched data



#	Sex	Tx Group	Cas9 VCN	Indels
68	F	sgRNA 1	20.5	0
69	M	sgRNA 1	22.73731	13
79	F	sgRNA 1	26.45161	0
82	M	sgRNA 1	17.47368421	31
136	F	sgRNA 2	36.64864865	0
135	F	sgRNA 2	41.64893617	0
138	M	sgRNA 2	51.98924731	0
139	M	sgRNA 2	18.57142857	0
70	M	sgRNA 3	19.52153	0
231	F	sgRNA 3	35	0
232	F	sgRNA 3	40.6372549	0
236	M	sgRNA 3	53.91534392	0
88	F	sgRNA 4	31.10048	0
89	F	sgRNA 4	22.435	0
90	M	sgRNA 4	4.041995	0

Appendices Table 6 2: Raw experimental data for Guide screening experiment.

Summary of individual animal data from in guide screening experiment. Each row corresponds to a single mouse (animal ID shown in column 1). Mice are supported by treatment group, then by numerical order. Treatment groups are indicated as Cas9 with candidate sgRNAs 1, 2,3 and 4. Measured outcomes include Cas9 vector copy number and indel frequency (%).

`

#	Tx Group	Cas9 VCN	GFP VCN	Indels	Transcript	Junctions correct	Junctions reverse	Across the cut site	FACS	IHC GFP% 10X used in graphs	IHC GFP% Mosaics*
167	Splice GFP donor only	0.00	40.53	0	0.00%	0.00%	0.00%	100.00%		0.86%	
170	Splice GFP donor only	0.00	18.90	0	0.00%	0.00%	0.00%	100.00%		0.98%	
199	Splice GFP donor only	0.00	14.69	0	0.00%	0.00%	0.00%	100.00%		0.51%	2.02%
168	Cas9 + Splice GFP	29.32	38.27	11.00%	38.06%	9.23%	6.40%	74.65%		72.66%	
187	Cas9 + Splice GFP			6.00%	25.57%	5.55%	4.38%	67.56%		30.97%	
188	Cas9 + Splice GFP	31.83	30.47	14.00%	39.06%	8.54%	8.15%	79.45%		62.95%	
195	Cas9 + Splice GFP	17.65	18.79	0.00%	15.50%	2.21%	2.22%	64.86%	25.00%	18.27%	
215	Cas9 + Splice GFP	17.02	15.46	0.00%	19.12%	2.49%	2.09%	70.51%	27.50%	19.94%	23.02%
165	2PA GFP donor only	0.00	50.46	0	0.00%	0.00%	0.00%	100.00%			
185	2PA GFP donor only	0.00	47.81	0	0.00%	0.00%	0.00%	100.00%	0.00%	1.23%	
214	2PA GFP donor only	0.00	6.56	0	0.00%	0.00%	0.00%	100.00%	0.76%	0.19%	0.26%
166	Cas9 + 2PA GFP	17.80	26.78	0.00%	19.21%	3.03%	2.58%	91.89%		18.81%	
171	Cas9 + 2PA GFP	23.49	16.45	5.00%	16.62%	5.57%	3.48%	92.87%		32.54%	
174	Cas9 + 2PA GFP	22.86	14.25	15.00%	35.92%	7.61%	4.65%	74.67%			
220	Cas9 + 2PA GFP	15.66	11.12	0.00%	21.01%	3.82%	2.27%	90.69%	17.70%		
223	Cas9 + 2PA GFP	30.99	26.97	18.00%	44.51%	3.82%	2.27%	72.62%	51.70%	42.55%	59.43%
204	Splice PAH donor only	0.00	64.06	0	0.00%	0.00%	0.00%	100.00%			
219	Splice PAH donor only	0.00	72.55	0	0.00%	0.00%	0.00%	100.00%			
194	Cas9 + Splice PAH	6.73	23.06	0.00%	1.60%	2.95%	1.63%	88.53%			
200	Cas9 + Splice PAH	23.24	68.60	0.00%	7.04%	6.43%	4.14%	82.16%			
205	Cas9 + Splice PAH	13.14	38.69	2.00%	7.64%	7.90%	4.16%	92.02%			
213	Cas9 + Splice PAH	24.72	75.58	16.00%	8.88%	6.57%	3.86%	67.77%			
196	2PA PAH donor only	0.00	65.79	0	0.00%	0.00%	0.00%	100.00%			
207	2PA PAH donor only	0.00	50.57	0	0.00%	0.00%	0.00%	100.00%			
221	2PA PAH donor only	0.00	51.49	0	0.00%	0.00%	0.00%	100.00%			
181	Cas9 + 2PA PAH	11.59	17.80	1.00%	5.07%	6.15%	3.48%	114.56%			
183	Cas9 + 2PA PAH	2.75	8.22	2.00%	1.91%	1.23%	1.36%	100.08%			
201	Cas9 + 2PA PAH	17.28	44.46	0.00%	4.38%	3.89%	3.14%	71.55%			
222	Cas9 + 2PA PAH	20.39	39.02	3.00%	3.59%	5.38%	2.65%	115.88%			

Appendices Table 6 3: Raw experimental data for Cas9 + donor vector-treated cohort.

Summary of individual animal data from *in vivo* genome editing experiments. Each row corresponds to a single mouse (animal ID shown in column 1). Mice are supported by treatment group, then by numerical order. Treatment groups are indicated as Cas9 with splice-donor GFP, Cas9 with 2PA-donor GFP, Cas9 with splice-donor PAH, or Cas9 with 2PA-donor PAH. Measured outcomes include vector copy number (VCN) for Cas9 and donor (GFP or PAH), indel frequency (%), proportion of hybrid transcript detected (%), proportion of correctly and reversely spliced junctions (%), proportion of integration events spanning the cut site, and transduction efficiency assessed by FACS (GFP+ hepatocytes, %) and immunohistochemistry (IHC, GFP+ hepatocytes, %)