Monogenic causes of severe fetal abnormalities leading to prenatal or perinatal lethality: Lessons from Neu-Laxova syndrome

Thesis

for the degree of

Doctor rerum naturalium (Dr. rer. nat.)

approved by the Faculty of Natural Sciences of Otto von Guericke University Magdeburg

by

(M.Sc., Fatima Abdelfattah)

born on 18. 11. 1980 in Aleppo

Examiner:

Prof. Dr. med. Martin Zenker

Prof. Dr. rer. nat. Peter Burfeind

submitted on: 19. 12. 2022
defended on: 15. 06. 2023

Dedication

To my family

To all people who want to live together in peace and harmony without greed and wars.

Your Fatima

Table of Contents

Abstract

Serine biosynthesis disorders (SBD) are a group of neurological and multisystem disorders caused by a defect in one of the three synthesizing enzymes of the L-serine biosynthesis pathway. They comprise a spectrum of very rare autosomal recessive inborn errors of metabolism with wide phenotypic variability. Neu-Laxova syndrome (NLS) represents the most severe expression, and it is characterized by multiple congenital anomalies and pre- or perinatal lethality. L-serine is a non-essential amino acid and plays manifold roles in metabolism, cell proliferation, growth, and differentiation due to its importance in the one-carbon unit metabolism and the availability of methyl groups. L-serine is also particularly important for development and function of the central nervous system (CNS), since it is the precursor of two neuroactive signaling molecules: glycine and D-serine. This study aimed to extend the mutation spectrum causing SBD, including the identification of novel mutations in already known genes as well as possible novel genes, and at a detailed phenotypic analysis in 15 unrelated families with severe types of SBD to delineate genotype-phenotypes correlations. This study revealed likely disease-causing variants in the *PHGDH* and *PSAT1* genes, several of which have not been reported previously. Phenotype analysis and a comprehensive review of the literature corroborated the evidence that SBD represent a continuum with varying degrees of phenotypic severity.

It has been assumed that the variability in the phenotype results from the degree of the residual enzyme activity. To check the validity of this assumption, and because PHGDH is the first enzyme in serine biosynthesis pathway, this study focused on functional studies of *PHGDH*. Modeling of *PHGDH* defects of both, NLS and non-lethal serine biosynthesis disorders (SBDNL), was done using genome editing tools. Knock-out *PHGDH* HEK293 cells (koHEK)

were established using CRISPR/Cas9 and these cells were later used to transiently express different mutants of *PHGDH in vitro*. Also, via gene targeting of the endogenous safe harbor locus *AAVS1* of these koHEK cells, several stable models of mutant *PHGDH* were established using TALEN system as another genome editing tool. The transient and stable mutant *PHGDH* cell lines were then used to quantify the residual activity of mutant PHGDH enzymes and demonstrate their effect on the proliferation of the cells. This approach proved useful to confirm the functional incapacity of various SBD-associated mutants and pointed at different modes of inactivation. The employed model systems could also partially demonstrate different levels of residual PHGDH activity correlating with disease severity.

Zusammenfassung

Serin-Biosynthese-Defekte (SBD) sind eine Gruppe von Multisystem-Erkrankungen mit prädominierenden neurologischen Störungen, welche durch einen Defekt in einem der drei Enzyme des L-Serin-Biosyntheseweges verursacht werden. Sie umfassen ein Spektrum sehr seltener autosomal-rezessiver angeborener Stoffwechselstörungen mit großer phänotypischer Variabilität. Das Neu-Laxova-Syndrom (NLS) stellt die schwerste Form dar und ist durch multiple angeborene Anomalien und prä- oder perinatale Letalität gekennzeichnet. L-Serin ist eine nicht essentielle Aminosäure und spielt durch seine Bedeutung im Stoffwechsel von C1-Einheiten und für die Verfügbarkeit von Methylgruppen eine vielfältige Rolle im Stoffwechsel, der Zellproliferation, dem Wachstum und der Differenzierung. Daneben ist L-Serin auch wichtig für die Entwicklung und Funktion des zentralen Nervensystems, da es die Vorstufe zweier neuroaktiver Signalmoleküle ist: Glycin und D-Serin.

Diese Doktorarbeit zielte darauf ab, das Mutationsspektrum bei SBD zu erweitern, einschließlich der Identifizierung neuer Mutationen in bereits bekannten Genen sowie möglicher neuer Gene, und vollzog eine detaillierte phänotypische Analyse in 15 nicht verwandten Familien mit schweren Formen von SBD, um Genotyp-Phänotyp-Korrelationen aufzuzeigen. Es wurden wahrscheinlich krankheitsverursachende Varianten in den Genen *PHGDH* und *PSAT1* identifiziert, von denen einige zuvor noch nicht bekannt waren. Die Phänotyp-Analyse und eine umfassende Überprüfung der Literatur bestätigen, dass SBD ein Kontinuum mit unterschiedlichem Grad an phänotypische Ausprägung darstellt. Es wird angenommen, dass die Variabilität des Phänotyps aus dem Grad der verbleibenden Enzymaktivität resultiert. Um die Gültigkeit dieser Annahme zu überprüfen und da PHGDH das erste Enzym im Serin-Biosyntheseweg ist, konzentrierte sich diese Doktorarbeit auf

funktionelle Studien von *PHGDH*. Die Modellierung von *PHGDH*-Defekten sowohl bei NLS- als auch bei nicht letalen Formen des Serinbiosynthese-Defekts (SBDNL) erfolgte unter Verwendung von Genom-Editing-Tools. Unter Verwendung von CRISPR/Cas9 wurden Knock-out-*PHGDH* HEK293 (koHEK)-Zellen etabliert und diese Zellen wurden später verwendet, um transient unterschiedliche Mutanten von *PHGDH in vitro* zu exprimieren. Darüber hinaus wurden durch Gen-Targeting des endogenen Safe-Harbor-Locus *AAVS1* dieser koHEK-Zellen mehrere stabile Expressions-Modelle der mutierten PHGDH unter Verwendung des TALEN-Systems als weiteres Genom-Editierungswerkzeug etabliert. Die transienten und stabilen mutanten *PHGDH*-Zelllinien wurden wiederum verwendet, um die Aktivität der verschiedenen *PHGDH* Mutationen zu quantifizieren und ihren Einfluss auf die Zellproliferation zu beschreiben. Dieser Ansatz erwies sich als geeignet, den funktionellen Defekt SBD-assoziierter *PHGDH*-Mutationen zu belegen und wies darauf hin, dass unterschiedliche Mechanismen der Inaktivierung vorliegen können. Die verwendeten Modellsysteme konnten teilweise unterschiedliche Grade von Enzymrestaktivität zeigen, die mit dem Schweregrad der Erkrankung korrelieren.

List of Figures

List of Tables

List of Abbreviations

ANOVA	Analysis of Variance
3-PG	3-phosphoglycerate
AAV	Adeno-associated virus
AAVS1	Adeno-associated virus integration site1
ACMG	American College of Medical Genetics
BrdU	5-bromo-2'-deoxyuridine
cDNA	Complementary DNA
CMV	Cytomegalovirus
CNS	Central nervous system
CRISPR	Clustered regularly interspaced short palindromic DNA repeats
crRNA	CRISPR RNAs
CSF	Cerebrospinal fluid
C_T	Threshold cycle
CVS DNA	Chroionic villus sampling DNA
DMEM	Dulbecco's Modified Eagle Medium
DSBs	Double strand breaks
dsDNA	Double-stranded DNA
E	Amplification efficiency
FBS	Fetal bovine serum
FFPE	Formalin-fixed paraffin-embedded tissue
FS	Frameshift
gDNA	Genomic DNA
GFP	Green fluorescent protein
gRNA	Guide RNA
HDR	Homology-directed repair
HEK293	Human embryonic Kidney 293
HR	Homologous recombination
indels	Insertions/deletions
IUGR	Intrauterine growth restriction
koHEK	Knock-out *PHGDH* HEK293 cells
LOVD	Leiden Open Variation Database
MEM	Minimum Essential Medium
MRI	Magnetic resonance imaging
MTS	Tetrazolium compound
NADPH	Nicotinamide adenine dinucleotide phosphate
NHEJ	Non-homologous end-joining
NLS	Neu-Laxova syndrome
nt	Nucleotide
OFP	Orange fluorescent protein
ORF	Open reading frame
PAM	Protospacer adjacent motif
PCR	Polymerase chain reaction

PGK	Phosphoglycerate kinase
PHGDH	3-phosphoglycerate dehydrogenase
PLP	pyridoxal 5'-phosphate
pre-crRNA	pre-CRISPR RNA
PSAT1	Phosphoserine aminotransferase
PSPH	Phosphoserine phosphatase
qPCR	Quantitative real-time PCR
R^2	Coefficient of determination
RNAi	RNA interference
RVD	Repeat variable di-residues
SBD	Serine biosynthesis defects
SBDNL	Non-lethal serine biosynthesis defects
SHMT	Serine hydroxymethyl transferase
SOB	Super optimal broth
SOC	Super optimal broth with catabolite repression
TALENs	Transcription activator-like effector nucleases
TALEs	Transcription activator-like effectors
T_m	Melting temperature
TOPO cloning	Topoisomerase based cloning
tracr RNA	Trans activating RNA
UPD	Uniparental disomy
UTR	Untranslated region
VUS	Variant of uncertain significance
WES	Whole exome sequencing
wtHEK	Wild type of human embryonic kidney293 cells
ZFNs	Zinc finger nucleases

1 Introduction

1.1 Serine biosynthesis pathway

L-serine is a non-essential amino acid in eukaryotic cells and is used for protein synthesis and in production of precursors for phosphoglycerides, glycerides, sphingolipids, phosphatidylserine, and methylenetetrahydrofolate (Murtas, Marcone, Sacchi, & Pollegioni, 2020). Morever, L-serine is the precursor of several other amino acids like glycine and D-serine, which represent the neuroactive signaling molecules of the neural N-methyl-D-aspartate receptors (Wang *et al.*, 2013). Serine can also play a role in the regulation of the redox status since serine is involved in the production of nicotinamide adenine dinucleotide phosphate (NADPH) (Ravez, Spillier, Marteau, Feron, & Frédérick, 2017). In summary, serine can provide nucleotide, lipid, amino acid, and cofactor building units, and thus it plays a versatile role in the metabolism and cell prolifiration (Mattaini, Sullivan, & Vander Heiden, 2016).

L-serine can be derived from diet, protein and glycine degradation by serine hydroxymethyl transferase (SHMT), but is also produced in considerable amounts by *de novo* biosynthesis (Furuya, 2008). It was previously described that *de novo* L-serine biosynthesis occurs by two different pathways, the non-phosphorylated pathway from 2-phosphoglycerate, which actually functions as a route for serine breakdown and the phosphorylated pathway from 3-phosphoglycerate (3-PG), which appears the sole true physiological pathway and for serine biosynthesis (Keith Snell, 1986). In humans, ~75% of the L-serine that appears during fasting comes from *de novo* serine biosynthesis (Kalhan & Hanson, 2012).

The phosphorylated *de novo* biosynthesis pathway starts from 3-PG generated during glycolysis. It involves three enzymes: 3-phosphoglycerate dehydrogenase (PHGDH, EC 1.1.1.95), phosphoserine aminotransferase (PSAT1, EC 2.6.1.52) and phosphoserine

phosphatase (PSPH, EC 3.1.3.3) (**Figure 1**), which are coordinately expressed in many tissues (brain, kidney, liver, skin, testis, and spleen).

In the CNS, owing to the low permeability of the blood-brain barrier and low transport of serine across it, the phosphorylated *de novo* biosynthesis pathway represents the main source of L-serine and is synthesized in astrocytes rather than neurons (Tabatabaie, Klomp, Berger, & de Koning, 2010). Therefore, a lot of effects and phenotypes regarding the nervous system were reported (as shown later in paragraph 1.2) owing to genetic defects in the three enzymes of this pathway (El-Hattab *et al.*, 2016; J. H. Yang *et al.*, 2010).

Upregulation of serine biosynthesis pathway, espically the increased expression of the first rate-limiting enzyme PHGDH, have been noticed in homogenates of neoplastic tissues relative to normal tissues (Locasale *et al.*, 2011; Possemato *et al.*, 2011; K. Snell & Weber, 1986). Furthermore, it was shown that this increased PHGDH activity might be a relevant driver for tumorigenesis (DeBerardinis, 2011; Mattaini *et al.*, 2016). It was also found that extracellular L-serine alone is not able to support cell proliferation. This has led to the hypothesis that the pathway is performing some function related to tumor growth other than supplying L-serine (G. A. Grant, 2018).

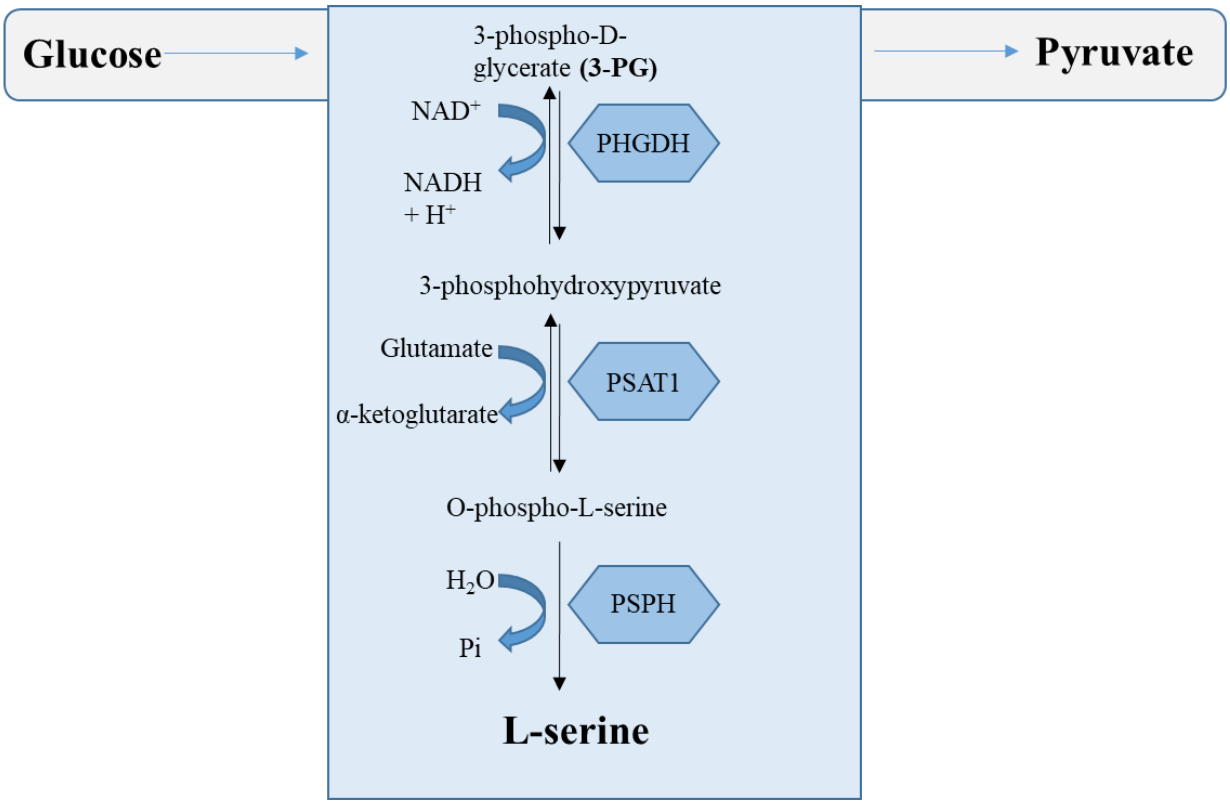


Figure 1: Schematic model of de novo serine biosynthesis pathway.

De novo biosynthesis of serine represents the main source to provide the organism with sufficient levels of L-serine. Glucose is converted to L-serine via glycolysis and the phosphorylated pathway. The first and limiting step of this pathway is the conversion of 3-PG to 3-phosphohydroxypyruvate by 3-phosphoglycerate dehydrogenase (PHGDH). This is followed by the conversion of 3-phosphohydroxypyruvate to o-phospho-L-serine by Phosphoserine aminotransferase (PSAT1). This reaction is accompanied by the transformation of glutamate to α-ketoglutarate. The final step in the serine biosynthesis pathway is catalyzed by phosphoserine phosphatase (PSPH), which gives rise to L-serine.

PHGDH represents the first enzyme in the serine biosynthesis pathway and is a cytosolic enzyme that reversible NAD^+-coupled oxidation of the glycolytic intermediate 3-PG to 3-phosphohydroxypyruvate (**Figure 1**). The human *PHGDH* coding gene consists of 12 exons and has been mapped to chromosome 1p12 by fluorescence *in situ* hybridization (Baek, Jun, Taub, & Kim, 2000). The nucleotide sequence of the human *PHGDH* gene was determined by Cho *et al*. (H. M. Cho, Jun, Bae, Ahn, & Kim, 2000). A human PHGDH-encoding transcript has been detected at high levels in the prostate, testis, ovary, brain, liver, kidney , and pancreas, whereas lower levels are present in the thymus, colon, and heart (Baek, Jun, Taub, & Kim,

2003) and it encodes a 533 amino acid protein (~57 kDa) and shares a 94,6% and 36,8% sequence identity with PHGDH from *R. norvegicus* and *M. tuberculosis*, respectively (Achouri, Rider, Schaftingen, & Robbi, 1997; Gregory A. Grant, 2018).

PSAT1 represents the second enzyme in the serine biosynthesis pathway and catalyzes the reversible glutamate-linked transamination of 3-phosphohydroxypyruvate to 3-phosphoserine, concomitantly producing α-ketoglutarate (**Figure 1**). The human *PSAT1* coding gene consists of 9 exons and has been mapped to chromosome 9q21.2. *PSAT1* encodes 370 amino acid protein (~40 kDa). The nucleotide sequence of the human *PSAT1* gene was determined by Baek *et al*. (Baek *et al.*, 2003). Compared to the human *PHGDH* encoding gene, the human *PSAT1* gene seems to exhibit a more restricted tissue-specific expression, since the transcript of *PSAT1* has been detected in the brain, liver, kidney, and pancreas (Baek *et al.*, 2003).

PSPH catalyzes the final and irreversible step of the serine biosynthesis pathway (Vincent *et al.*, 2015) (**Figure 1**). The human *PSPH* coding gene consists of 8 exons and has been mapped to chromosome 7p11.2 by (Veiga-da-Cunha *et al.*, 2004). Human *PSPH* was cloned by Collet *et al*. (Collet, Gerin, Rider, Veiga-da-Cunha, & Van Schaftingen, 1997) and it encodes a 225 amino acid protein (~25 kDa) and is expressed in a wide range of tissues, but its activity differs with the highest activity being detected in the hippocampus in the human brain and the lowest in the cortex (Veeranna & Shetty, 1990).

1.2 Serine biosynthesis disorders (SBD)

SBD constitutes a clinically and genetically heterogeneous group of autosomal recessive diseases caused by pathogenic changes in genes encoding enzymes of the L-serine biosynthesis pathway: phosphoglycerate dehydrogenase (*PHGDH*, MIM *606879), phosphoserine

aminotransferase (*PSAT1*, MIM *610936), and phosphoserine phosphatase (*PSPH*, MIM *172480) (Acuna-Hidalgo *et al.*, 2014; Hart *et al.*, 2007; Klomp *et al.*, 2000; Shaheen *et al.*, 2014; Veiga-da-Cunha *et al.*, 2004). The phenotype ranges from non-specific developmental delay (Tabatabaie *et al.*, 2011) to the severe lethal disease known as Neu-Laxova syndrome (NLS; MIM# PS256520) presenting with congenital neurological dysfunction and psychomotor retardation of variable degree (Abdelfattah *et al.*, 2020; El-Hattab *et al.*, 2016).

1.2.1 PHGDH deficiency

Phosphoglycerate dehydrogenase deficiency (MIM# 601815) was first reported by Jaeken *et al.* (Jaeken *et al.*, 1996), when they described the clinical and biochemical features of two brothers whose parents were first cousins. Both patients presented congenital microcephaly, profound psychomotor retardation, hypertonia, epilepsy, growth retardation, and hypogonadism. They exhibited markedly low serine levels in both, plasma, and cerebrospinal fluid (CSF), in a fasting state. In fibroblasts of both patients, a decreased activity of the first enzyme of serine biosynthesis pathway was demonstrated, with a residual activity of PHGDH ranging between 13% and 22% of the mean control value, whereas the other enzymes of the pathway, PSAT1 and PSPH, showed normal activity. Despite the severe neurological and behavioral deficits in both patients, one of them responded to the oral serine supplementation with cessation of convulsions after one week of oral treatment. Therefore, it has later been suggested that PHGDH deficiency is a potentially treatable inborn error of metabolism (de Koning *et al.*, 1998; Jaeken *et al.*, 1996; Pineda *et al.*, 2000).

The molecular basis of PHGDH deficiency has been characterized in more detail by Klomp *et al.* (Klomp *et al.*, 2000), since they recognized the first two disease-associated mutations (c.1273G>A; p.Val425Met, c.1468G>A; p.Val490Met) in the carboxyterminal part of

PHGDH, and *in vitro* expression of these mutant proteins resulted in significant reduction of PHGDH enzyme activities (Klomp *et al.*, 2000). After the initial report, several other patients with PHGDH deficiency have been reported and the corresponding *PHGDH* variants have been recognized (Ali, Dhahouri, Almesmari, Fathalla, & Jasmi, 2021; Benke *et al.*, 2017; Brassier *et al.*, 2016; Coşkun *et al.*, 2009; de Koning *et al.*, 1998; El-Hattab *et al.*, 2016; Hausler, Jaeken, Monch, & Ramaekers, 2001; Klomp *et al.*, 2000; Kraoua *et al.*, 2013; Meneret *et al.*, 2012; Pind *et al.*, 2002; Pineda *et al.*, 2000; Poli *et al.*, 2017; Tabatabaie *et al.*, 2009; Tabatabaie *et al.*, 2011). Most of these patients presented with congenital microcephaly, intrauterine growth restriction (IUGR), severe psychomotor retardation and intractable seizures. The symptoms of the majority of these patients started prenatally or perinatally except for two reported milder expressions of the disease: the first one with mild psychomotor retardation and a juvenile onset of seizures was reported in two children of healthy nonconsanguineous Dutch parents (Tabatabaie *et al.*, 2011). The second one was reported with a 31-year old man with a polyneuropathy and progressive motor disability in the upper and lower limbs (Meneret *et al.*, 2012). Studies in various patient fibroblast samples showed decreased PHGDH enzyme activity compared to controls (Benke *et al.*, 2017; Jaeken *et al.*, 1996; Pind *et al.*, 2002; Pineda *et al.*, 2000; Tabatabaie *et al.*, 2009) . In addition to abnormal serine and glycine levels in plasma and CSF, metabolomic profiling in some patients revealed low glycerophospholipids including low phosphatidylcholine (Glinton *et al.*, 2018), further supporting that PHGDH plays an important role in CNS development. However, severity of symptoms does not always correlate with plasma or CSF serine levels (van der Crabben *et al.*, 2013).

1.2.2 PSAT1 deficiency

Phosphoserine aminotransferase deficiency (MIM# 610992) was first reported by Hart *et al*. (Hart *et al.*, 2007) in two siblings born to nonconsanguineous British parents, who showed low levels of serine and glycine in plasma and CSF. The index patient presented with acquired microcephaly, psychomotor retardation, muscular hypertonia, and uncontrollable seizures. Cranial imaging showed generalized cerebral atrophy, a hypoplastic cerebellar vermis, and abnormal white matter development (Hart *et al.*, 2007). This patient died at the age of 7 months. The younger sister showed normal fetal development by prenatal monitoring, but CSF and plasma samples revealed low concentrations of serine and glycine after birth, and the supplementation with serine (500 mg/kg/d) and glycine (200 mg/kg/d) was initiated within the first day of life. Brain ultrasound and magnetic resonance imaging (MRI) at age 4 months showed no abnormalities. The mutation analysis of both siblings revealed compound heterozygous mutations in the *PSAT1* gene (c.299A>C, c.107delG) (Hart *et al.*, 2007). The effect of the missense mutation was tested in the bacterial expression system and the recombinant mutated PSAT1 showed low solubility and low enzyme activity (Hart *et al.*, 2007). Following the initial report, two additional cases of PSAT1 deficiency have been reported before this study was started (Brassier *et al.*, 2016; Glinton *et al.*, 2018). All these patients showed IUGR, progressive microcephaly, hypoplasia and several abnormalities in CNS, intellectual disability, severe psychomotor retardation, failure to thrive, severe feeding difficulties, and joint contractures. Cataracts and mild ichthyosis have been reported in one patient, who had this mutation (c.129T>G; p.Ser43Arg) and this patient was of Turkish origin (Brassier *et al.*, 2016).

In 2020, Sirr and his colleagues measured the functional impact of most published *PSAT1* variants by a yeast-based growth assay and their results demonstrated that all pathogenic alleles tested exhibited significantly reduced growth in the absence of serine (relative to a strain homozygous for the wild type *yPSAT1*) (Sirr *et al.*, 2020). In addition, they found that the genotypes that produced growth values of 79%-82% of wild-type growth rates in the yeast assay corresponded to more mildly affected individuals who survived past 7 months of age like the patients reported by Brassier *et al*. and Glinton *et al*. (Brassier *et al.*, 2016; Glinton *et al.*, 2018).

1.2.3 PSPH deficiency

Phosphoserine phosphatase deficiency (MIM# 614023) was reported for the first time in a patient with Williams syndrome by Jaeken *et al*. (Jaeken *et al.*, 1997). This patient showed decreased serine levels in plasma and CSF, moderate psychomotor retardation, early feeding difficulties and facial dysmorphism suggestive of Williams syndrome including puffy eyelids, wide mouth, broad forehead and micrognathia. Further investigation showed that PSPH activity in lymphoblasts and fibroblasts of this patient reduced to 25% of normal values, whereas oral serine supplementation normalized the plasma and CSF serine levels and showed some clinical improvement (Jaeken *et al.*, 1997). In 2004, the genetic defect responsible for this case was identified and two mutations in *PSPH* gene (c.94G>A; p.Asp32Asn, c.155T>C; p.Met52Thr) were recognized (Veiga-da-Cunha *et al.*, 2004). Both mutations were studied on the basis of the human recombinant enzyme expressed in *Escherichia coli*, where (c.155T>C; p.Met52Thr) almost abolished the enzymatic activity, whereas the (c.94G>A; p.Asp32Asn) mutation caused a 50% reduction in enzyme activity (Veiga-da-Cunha *et al.*, 2004).

The second report described seven individuals from a consanguineous Pakistani family. All the affected individuals showed moderate to profound intellectual disability, hypertonia,

microcephaly, and seizures. MRI identified brain atrophy in one individual and all of them showed low serine in plasma and CSF. A homozygous missense mutation (c.103G>A, p.Ala35Thr) was recognized in *PSPH* in this family (Vincent *et al.*, 2015).

The third report described a novel presentation of PSPH deficiency in an adult with progressive myeloneuropathy and distal contractures of the upper extremities. This patient had no motor abnormalities throughout the childhood, but at age 19, she began to complain of increasing lower extremity stiffness. Gross and fine motor skills subsequently degraded, and her gait became progressively abnormal. Stiffness progressed to involve her upper extremities at age 32 and she developed contractures of her hands and fingers (Byers *et al.*, 2016). Molecular testing revealed two novel mutations in *PSPH* and serine supplementation led to improvement of the neuropathic pain in this patient (Byers *et al.*, 2016). This case expanded the phenotypic spectrum of PSPH deficiency.

1.2.4 Neu-Laxova syndrome

Neu-Laxova syndrome (NLS; MIM# PS256520) is an autosomal recessive lethal multiple malformation syndrome. It was described for the first time by Neu and Laxova in 1971 and 1972, respectively (Laxova, Ohara, & Timothy, 1972; Neu, Kajii, Gardner, & Nagyfy, 1971). Following the first descriptions less than 100 cases of NLS have been reported to date (Acuna-Hidalgo *et al.*, 2014; Bourque *et al.*, 2019; Cavole *et al.*, 2020; Coto-Puckett *et al.*, 2010; El-Hattab *et al.*, 2016; Manning, Cunniff, Colby, El-Sayed, & Hoyme, 2004; Mattos *et al.*, 2015; Ni *et al.*, 2019; Shaheen *et al.*, 2014). The clinical hallmarks of this disorder include severe IUGR, microcephaly, cutaneous abnormalities, and distinctive craniofacial features including sloping forehead, ocular hypertelorism, prominent eyes, ectropion, flat nose, round gaping mouth, micrognathia, short neck, and low-set malformed ears. Variable CNS abnormalities

have been described including microencephaly, lissencephaly, hypoplastic cerebellum, agenesis, or dysgenesis of the corpus callosum, and Dandy-Walker malformation (Badakali, Badakali, & Dombale, 2012; Coto-Puckett *et al.*, 2010; Ostrovskaya & Lazjuk, 1988). The spectrum of skin abnormalities comprises varying degrees of ichthyosis (Curry, 1982), edema of the hands and feet, and excessive fatty tissue under the epidermis (Karimi-Nejad, Khajavi, Gharavi, & Karimi-Nejad, 1987). Joint contractures are common, and pterygia may be present, most likely reflecting fetal akinesia (Curry, 1982; Ejeckam, Wadhwa, Williams, & Lacson, 1986; Fitch, Resch, & Rochon, 1982). Additional congenital malformations including cleft lip and palate (Coto-Puckett *et al.*, 2010; Rouzbahani, 1995), spina bifida (Manning *et al.*, 2004; Naveed, Manjunath, & Sreenivas, 1990), genitourinary anomalies (cryptorchidism, hypoplastic genitalia, renal agenesis) (Naveed *et al.*, 1990; Shivarajan, Suresh, Jagadeesh, Lata, & Bhat, 2003), pulmonary and gastric hypoplasia (Manning *et al.*, 2004) have occasionally been reported. Although pre- or perinatal lethality is a hallmark of this condition, survival from a few weeks to several months has been observed in some cases (Carder, Fitzpatrick, & Weston, 2003; Coto-Puckett *et al.*, 2010; El-Hattab *et al.*, 2016; Horn, Muller, Thiele, & Kunze, 1997; Ugras, Kocak, & Ozcan, 2006), presumably representing milder variants of the syndrome.

In 2014, NLS was found to represent the severe end of SBD by Shaheen and colleagues. They discovered two mutations in *PHGDH* by a combination of autozygosity mapping and exome sequencing in three patients with NLS from three unrelated consanguineous Saudi families (Shaheen *et al.*, 2014). In parallel and the same year, Acuna-Hidalgo *et al*. reported a cohort of 12 unrelated families affected by NLS. This report presented evidence that NLS is a genetically heterogeneous syndrome since four mutations in *PSAT1* were discovered in six families, one mutation in *PSPH* in one family, and five mutations in *PHGDH* in three unrelated families. Two families were reported negative for a mutation in the three genes of the serine biosynthesis

pathway (Acuna-Hidalgo *et al.*, 2014), thus suggesting the existence of one or more NLS-causing genes. Following these, both reports, several mutations in either *PHGDH* or *PSAT1* were reported in seven case reports on NLS patients (Bourque *et al.*, 2019; Cavole *et al.*, 2020; Kapoor *et al.*, 2021; Mattos *et al.*, 2015; Ni *et al.*, 2019; Takeichi *et al.*, 2018). Excluding this work 12 different variants in the *PHGDH* gene, 5 in the *PSAT1* gene, and one frameshift (FS) variation in *PSPH* have been identified in association with NLS until the day.

There is an obvious wide variability in the phenotypes associated with SBD. It was therefore proposed that the individual residual enzyme activity of mutant proteins might be the major cause of this variability (El-Hattab *et al.*, 2016). Using the PSAT1 yeast-based growth assay mentioned above (1.2.2) Sirr *et al*. found that genotypes that produced growth values of 51-69% of wild-type yPSAT1 activity in the yeast assay corresponded to the PSAT1 genotypes of affected individuals that were either stillborn or who died before the age of 7 months as patients of (Acuna-Hidalgo *et al.*, 2014; Hart *et al.*, 2007). The significant role of PHGDH was previously demonstrated in *PHGDH* knock-out mice, which displayed embryonic lethality, severe neurodevelopmental defects, and overall growth retardation (Furuya *et al.*, 2008; Yoshida *et al.*, 2004), but no functional PHGDH assay had been performed on mutations associated with NLS until the initiation of this work. Notably, even in the most severe expression of SBD, namely NLS, most mutations are missense, and patients with biallelic variants predicting a complete loss of function of PHGDH and PSAT1 are lacking. This suggests that NLS is associated with some residual activity of the mutant enzyme and that complete loss of activity might be lethal in the embryonic or early fetal period (Acuna-Hidalgo *et al.*, 2014).

1.3 Genome editing

Genome editing has revolutionized a wide range of fields, including the generation of transgenic animals, model development for diseases, drug development, and functional analysis of the genes in many organisms, where it has enhanced the ability of researchers to explain the contribution of genetics to disease by defining the relationship between a gene, its mRNA and the function of protein products. Through the inhibition of gene function or alteration of the sequence or expression of a gene, genome editing plays a key role and provides a powerful tool for studying the basis of disease (Li *et al.*, 2020).

1.3.1 Historical overview of genome editing and DNA repair mechanisms

The observations that exogenous DNA can be taken up by bacteria or yeast and be integrated randomly into the genome, is underlying the glow of the idea of genetic modification (Scherer & Davis, 1979). The first trial to target a specific gene was done in mouse embryonic stem cells using homologous recombination (HR) to generate knock-in and knock-out mutations (M. R. Capecchi, 1989). **Figure 2** shows the advancements in genome editing tools.

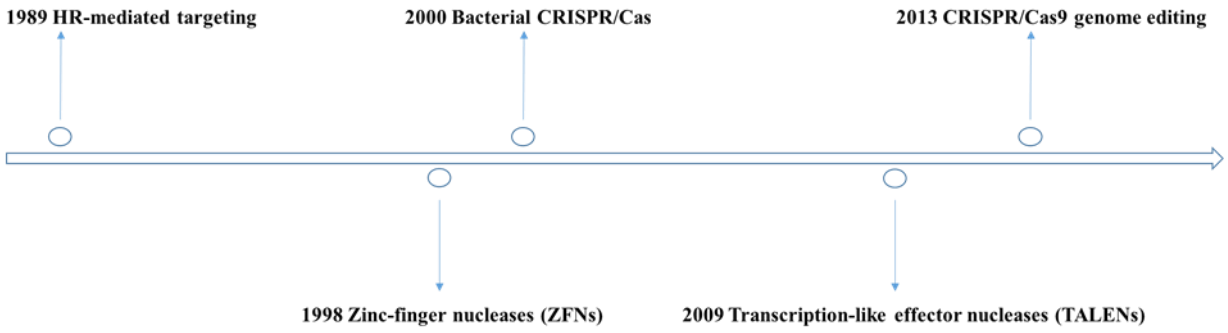


Figure 2: Advancements in genome editing tools.

HR-mediated targeting was the first trial to target a specific gene in 1989. In 1998, Zinc-finger nucleases (ZFNs) were engineered to induce the desired double-strand break at a particular DNA target site as an early development stage of genome editing. In 2000, the bacterial CRISPR/Cas system was discovered, which was harnessed later in

genome editing. In 2009, transcription-like effector nucleases (TALENs) were first described and utilized in genome editing. In 2013, the natural bacterial immune system was harnessed to modify DNA in a programmable manner which led to the development of CRISPR/Cas as a genome editing tool.

Since HR, which is naturally used by cells to repair double-strand breaks (DSBs) before mitosis and leads to new recombinations during meiosis, occurs infrequently in mammalian cells, the initial targeting frequencies were low ranging from 1 in 10^6 to 10^7 cells (M. Capecchi, 1990). HR technology was used as a tool to alter the genome for many years and was limited to certain organisms because of the need for complex targeting and selection constructs in addition to its low targeting frequencies. It was reported that the induction of DSBs into the genome can highly efficiently stimulate genome editing through HR mechanisms at the breakpoint and increase recombination rates significantly (Choulika, Perrin, Dujon, & Nicolas, 1995; Rouet, Smih, & Jasin, 1994). To improve editing rates of HR, customized and programmable engineered endonucleases, came into use in the late 1990s (Bibikova, Beumer, Trautman, & Carroll, 2003; Bibikova *et al.*, 2001). They represent more effective genome editing tools by employing DSB-mediated repair and included zinc finger nucleases (ZFNs) (Y. G. Kim, Cha, & Chandrasegaran, 1996) as well as transcription activator-like effector nucleases (TALENs) (Moscou & Bogdanove, 2009).

ZFNs and TALENs are artificial fusion proteins, and both have a modular structure that is composed of an engineered sequence-specific DNA-binding domain fused to a non-sequence-specific nuclease domain derived from the FokI restriction enzyme. Targeting these enzymes to given locus results in DNA cleavage and forces the cell to undergo the cellular DNA repair mechanisms. Two different DSB repair pathways can be taken place to achieve different genome modifications. If no donor DNA is present, non-homologous end-joining (NHEJ) causes by rejoining of the ends and owing to its error-prone nature, this usually leads to

nucleotide insertions and deletions (indels), which can result in FS mutations if they occur in the coding region of a gene, thereby effectively creating a knock-out gene. Alternatively, when an exogenous DNA template flanked by homologous sequences around the DSB site exists, then the homologous sequences are copied faithfully enabling precise editing and the cell employs a homology-directed repair (HDR) pathway in this case to repair the DSB (Rouet *et al.*, 1994; Shrivastav, De Haro, & Nickoloff, 2008).

The use of both systems is limited by the need to engineer specific proteins to each target site, which has been a bottleneck for many research labs (Joung & Sander, 2013; Urnov, Rebar, Holmes, Zhang, & Gregory, 2010). By 2013, the basic mechanism of CRISPR/Cas9 derived from *Streptococcus pyogenes* was clarified (Jinek *et al.*, 2012) and the latest tool for genome editing appeared to the light. This technology is based on RNA-guided engineered Cas9 nucleases, which have dramatically increased the ability to edit the genome of many organisms due to their simplicity, efficiency, and versatility (Bassett & Liu, 2014; Cong *et al.*, 2013; Dow, 2015; Hryhorowicz, Lipiński, Zeyland, & Słomski, 2017; Mali *et al.*, 2013; Seruggia & Montoliu, 2014; Torres-Ruiz & Rodriguez-Perales, 2017; L. Yang, Yang, Byrne, Pan, & Church, 2014). **Figure 3** shows the genome editing tools and mechanisms for DSB repair with endogenous DNA.

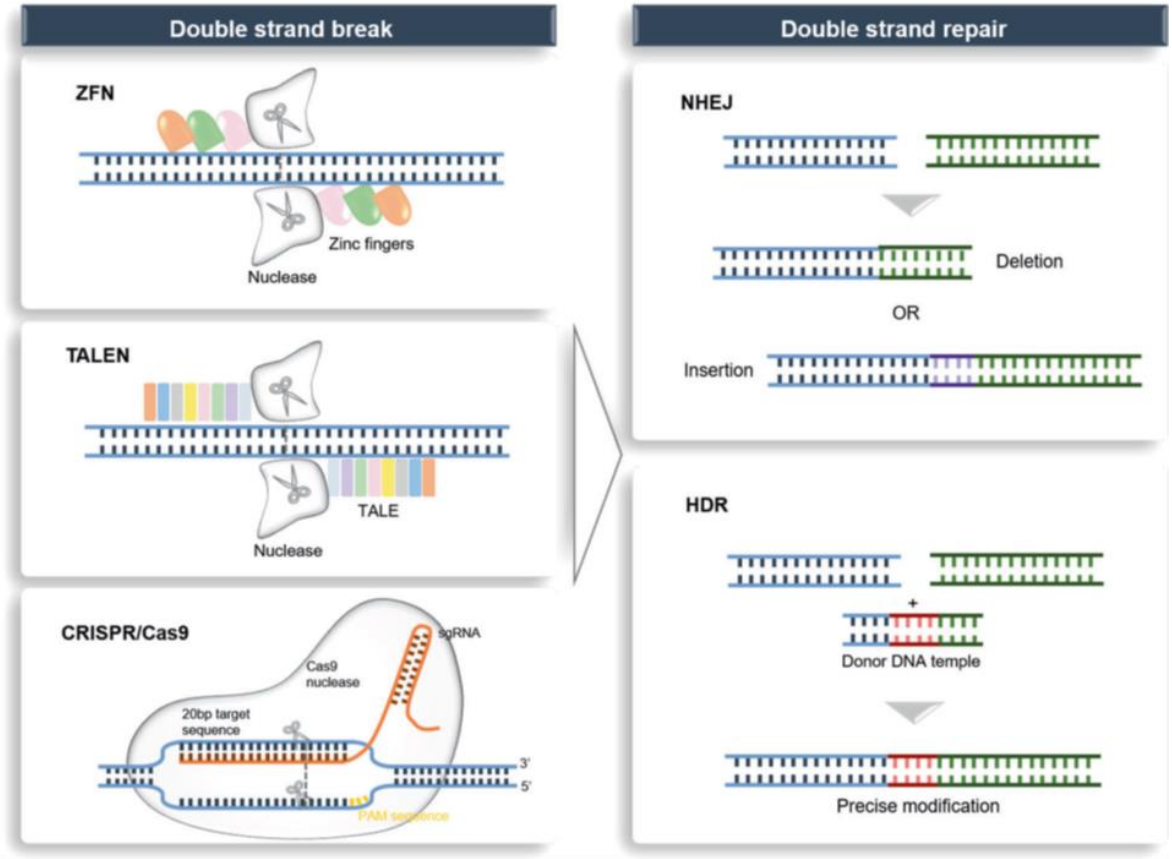


Figure 3: Genome editing tools and mechanisms for DSB repair with endogenous DNA.

Genome editing site-specific nucleases (ZFNs, TALENs, and CRISPR/Cas9) induce DSB at a specific site. DSB can be repaired either by non-homologous end joining (NHEJ) or homology-directed repair (HDR) in the presence of a donor template. Repair by NHEJ usually results in the insertion (purple) or deletion (green) of random base pairs, causing gene knock-out by disruption. HR with a donor DNA template can be exploited to modify a gene by introducing precise nucleotide substitutions. This figure is reprinted from Li et al. (Li et al., 2020). The permission to reuse this figure is under the terms of the Creative Commons CC BY license.

As shown in **Figure 3**, the three types of nucleases share the same mechanism of action, in that they cleave DNA in a site-specific manner, which in turn triggers an endogenous DNA repair system that results in genome editing and modification, but all of them are limited by the complexity of protein design and targeting efficiency that depended on selectable markers (Torres-Ruiz & Rodriguez-Perales, 2017). Comparing these genome editing tools to the systems used to interrogate gene function such as RNA interference (RNAi) which was used to silence a gene for over a decade, one could consider these three types of nucleases as the gold

standard for investigating gene function through their ability to disrupt the targeted gene and enable the complete loss of function to mimic and insert the desired genetic mutation which was difficult to achieve by RNAi (Barrangou *et al.*, 2015; Torres-Ruiz & Rodriguez-Perales, 2017).

1.3.1.1 TALEN overview

TALENs proteins were first identified in bacterial plant pathogens of the genus *Xanthomonas spp.*. Upon infection, the bacteria inject a number of Transcription activator-like effectors (TALEs) proteins via their type III secretion system, which can enter the nucleus of the plant cell and bind plant promoter sequences via a central domain of tandem repeats, and activate transcription of plant genes facilitating the bacterial infection (Boch & Bonas, 2010). TALENs are modular and artificial chemeric protein that consist of fusion a non specific FokI restriction endonuclease domain to a designed DNA recognition binding domain that directs the non-specific FokI endonuclease to the target site (Christian *et al.*, 2010) (**Figure 4**). This DNA binding domain consists of highly conserved 12 to 26 TALE repeat domains, which are made of repeating sequence of 33 to 35 amino acid arranged in tandem. The sequence of the amino acid of each TALE repeat domain is similar except for two-hypervariable amino acids located at positions 12 and 13 of the repeat called repeat variable di-residues (RVD) and this in turn confers the specificity of the individual TALE repeat in the array to recognize and bind to a single specific nucleotide of DNA (Boch *et al.*, 2009). It should be mentioned that TALENs must dimerize to create a DSB, but the dimer interface does not promote dimerization effectively (Cermak *et al.*, 2011). For this reason TALEN module is designed in pairs to bind opposing DNA target loci between the two binding sites when planning for genome editing (Bogdanove & Voytas, 2011).

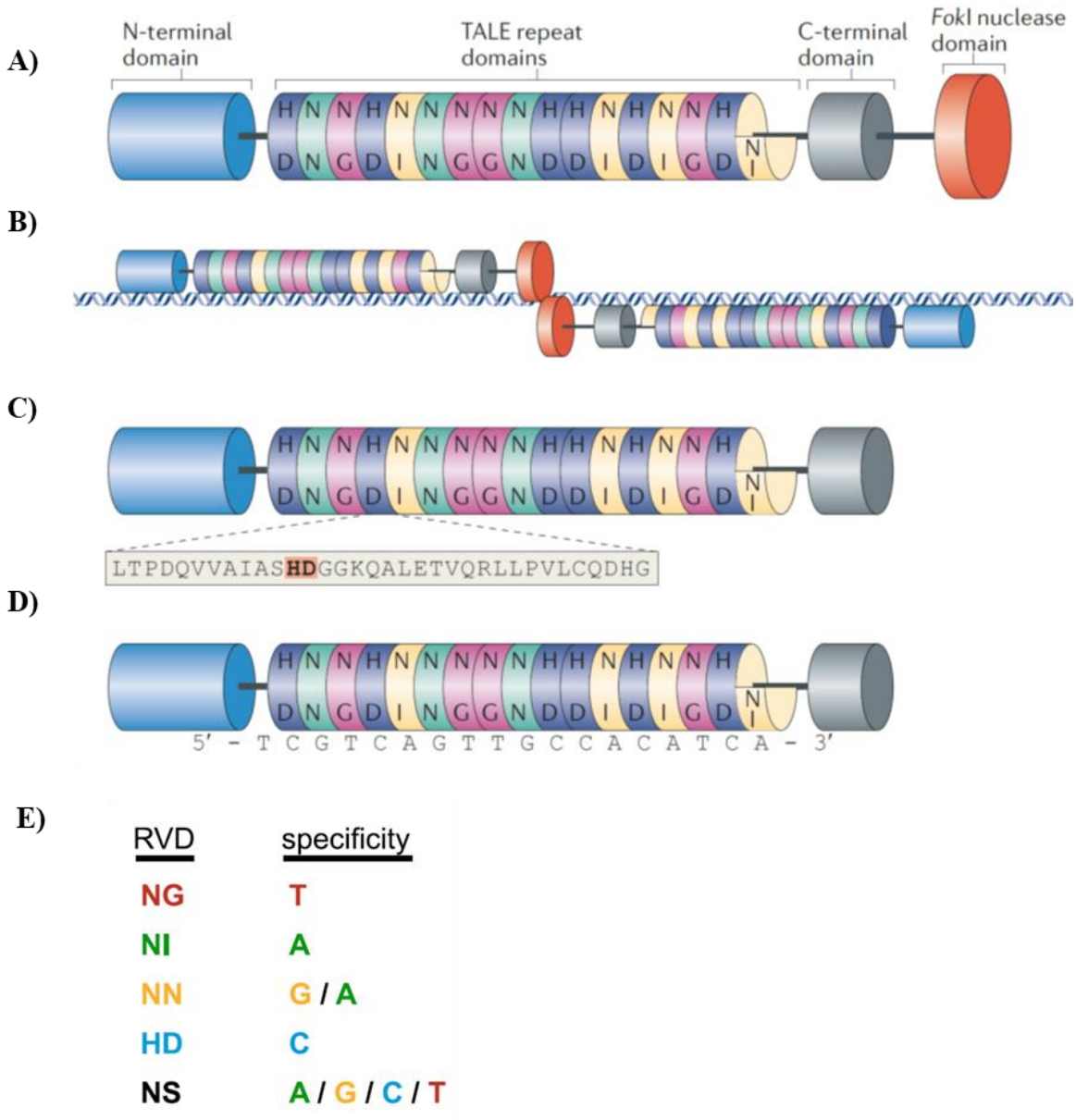


Figure 4: Overwiew of TALENs and TALE repeat arrays.

A) *Schematic diagram of TALEN. TALE repeats are shown as colored discs with a final carboxy-terminal truncated half repeat. Letters inside each repeat represent the two-hypervariable residues. TALE-derived N-terminal and C-terminal domains that are required for DNA binding are indicated. The nonspecific nuclease domain from the FokI endonuclease is shown in red.* ***B)*** *TALENs bind and cleave as dimers on a target DNA site. Note that the TALE-derived N-terminal and C-terminal domains flanking the repeats may also contact the DNA. Cleavage by the FokI nuclease domains occurs in the spacer sequence that lies between the two regions of the DNA bound by the two TALEN monomers.* ***C)*** *Schematic diagram of a TALE-derived DNA binding domain. The amino acid sequence of a single TALE repeat is expanded with the two-hypervariable residues highlighted in orange and bold letters.* ***D)*** *TALE-derived DNA-binding domain aligned with its target DNA sequence. Note that the repeat domains bind to single bases in the target sequence according to the TALE code. Also, note the presence of a 5 prime thymine preceding the first base bound by a TALE repeat.* ***E)*** *Shows the five most used repeat variable di-residues (RVD) and the nucleotides they specify. The figure with the respective legend is adapted from Becker,*

Joung, and their colleagues (Becker & Boch, 2021; Joung & Sander, 2013). The permission to reuse and adapt this figure is under the terms of the Creative Commons CC BY license.

TALENs have an advantage over ZFNs, because they are more flexible, and their DNA binding motifs can target a wider range of sequences. TALENs are useful for high-throughput gene editing, since the targeting rate is about 99% compared to ZFNs which have less targeting rate about 24%. Morover, the average mutation rate with TALENs is about 20% compared to ZFNs, which have a lower average mutation rate of about 10% (H. Kim & Kim, 2014; Reyon *et al.*, 2012). At the same time TALENs are expensive to produce, in adddition to the difficulties to assmebly TALEN constructs, and optimizing the custom proteins make their manipulation somewhat problematic in some cases. Another limitation to use TALENs seems in the design of TALENs, since a thymine at the 5 prime end of the target sequence is required (Mak, Bradley, Cernadas, Bogdanove, & Stoddard, 2012), in addition to the fact that TALENs cannot cleave target DNA that contains methylated cytosines (Bultmann *et al.*, 2012).

1.3.1.2 CRISPR/Cas9 overview and evolution of CRISPR

Clustered regularly interspaced short palindromic DNA repeats, which are known later as CRISPR loci, were first discovered in *Escherichia coli* in 1987 (Ishino, Shinagawa, Makino, Amemura, & Nakata, 1987). While studying the *IAP* gene, Ishino and colleagues reported a set of 29 nucleotides (nt) repeats downstream of the *IAP* gene and take the form of tandem repeats, since these 29 nt repeats were interspaced by five intervening 32 nt non-repetitive sequences. By 2000, additional studies had clarified that these clustered DNA repeats are present in the genomes of multiple different prokaryotic species since they are present in more than 40% of bacteria genomes and 90% of archaea genomes (F. J. Mojica, Díez-Villaseñor, Soria, & Juez, 2000). In 2002, it was discovered that these CRISPR loci are consistently located adjacent to the well-conserved *CAS* genes, which encode Cas endonucleases and that indicated that

CRISPR loci and *CAS* gene have a biological functional relationship (Jansen, Embden, Gaastra, & Schouls, 2002). In 2005, it was confirmed that these non-repetitive sequences in the CRISPR array, which are called spacers or protospacer, are derived from viral and foreign genetic sources (Bolotin, Quinquis, Sorokin, & Ehrlich, 2005; F. J. Mojica, Díez-Villaseñor, García-Martínez, & Soria, 2005). Later in 2007, it was reported that *Streptococcus thermophilus* can gain resistance against phages by integrating these spacers of the infectious phage into the CRISPR array and it was concluded that CRISPR/Cas system functions as an adaptive immune defense against the pathogen DNA (Barrangou *et al.*, 2007). Soon afterward, the mechanism for CRISPR/Cas-mediated immunity through phage DNA degradation was revealed (Brouns *et al.*, 2008; Deltcheva *et al.*, 2011; Marraffini & Sontheimer, 2008). To date, at least eleven different CRISPR/Cas systems have been discovered. These different systems have been grouped into three major types (I-III). The type II system from *Streptococcus thermophilus* is the best characterized one and it is known as CRISPR/Cas9 (Hsu, Lander, & Zhang, 2014; Makarova *et al.*, 2011). When a prokaryote is infected, a fragment of the sequence of the bacteriophage is inserted as a spacer into the *CRISPR* loci and this step is called adaptation (**Figure 5**). The *CRISPR* loci are then transcribed and resulting pre-CRISPR RNA (pre-crRNA). A trans-activating RNA (tracrRNA) and *Cas9* are transcribed separately. The tracrRNA first hybridizes to the short palindromic repeat and it triggers the processing of pre-crRNA by a Cas9/RNaseIII complex to generate mature CRISPR RNAs (crRNA). The resulting crRNA comprises Cas9 and tracrRNA a crRNA: tracrRNA: Cas9 complex and this step is called biogenesis (**Figure 5**). In this complex crRNA will work as a guide RNA (gRNA) and if any pathogen genome has a complementary nucleotide sequence to the crRNA sequence, then the nuclease Cas9 will be guided to the foreign DNA and introduce a DSB inducing the degradation of the phage DNA and this step is called as interference (**Figure 5**).

It is worth mentioning that the recognition of target DNA by Cas9 is triggered by the presence of a short 3 base pair sequence known as a protospacer adjacent motif (PAM), which is located adjacent upstream of protospacer (F. J. M. Mojica, Díez-Villaseñor, García-Martínez, & Almendros, 2009). PAMs are fundamental for Cas-mediated cleavage and targeting since they allow Cas9 to identify the bacterial DNA from the invading DNA (Mali *et al.*, 2013).

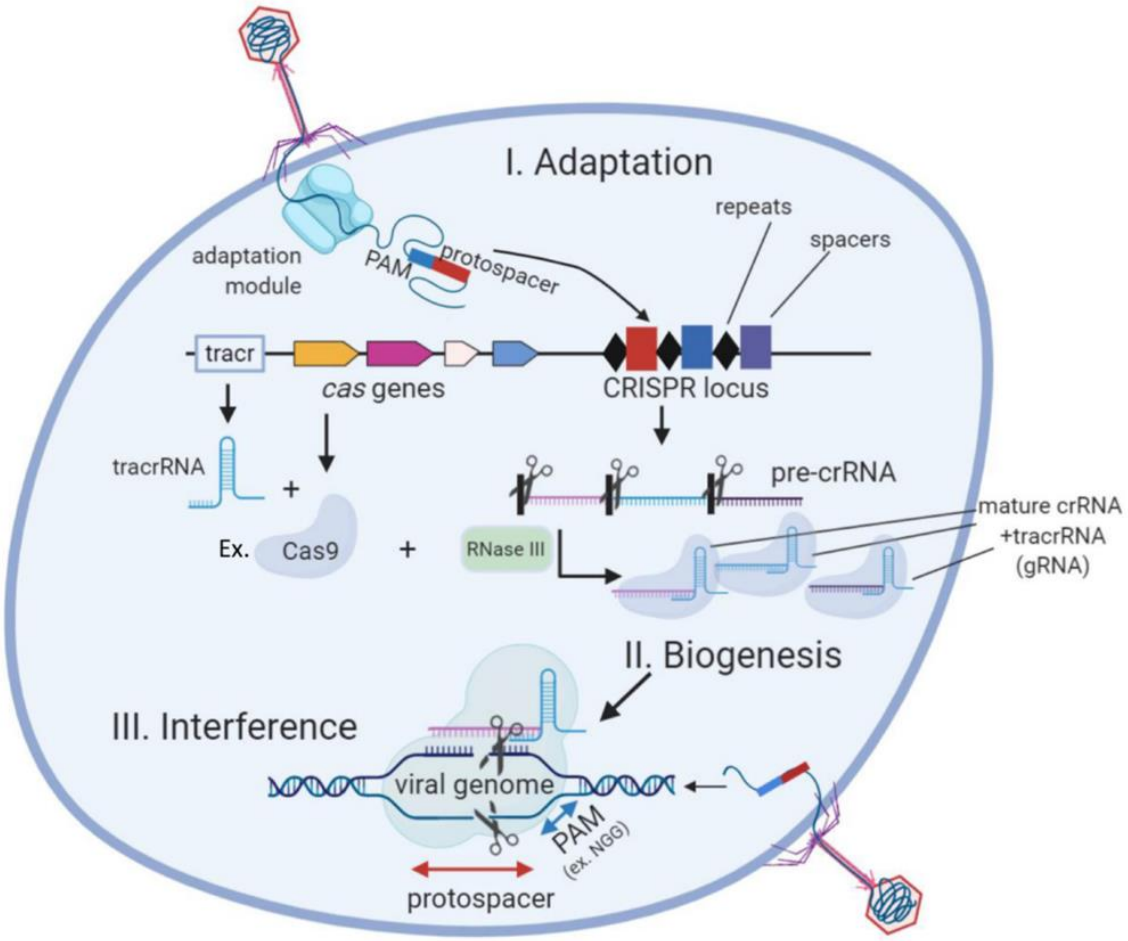


Figure 5: The three stages of the CRISPR/Cas9 (type II) bacterial adaptive immune system.

During CRISPR adaptation, the injection of phage DNA into bacterial cells activates the adaptation module proteins which excise spacer-sized fragments of phage DNA for incorporation into CRISPR loci. During CRISPR RNA biogenesis, CRISPR loci are transcribed, and resulting pre-crRNA, which is processed by a Cas9/RNaseIII complex at repeat sequences to generate mature crRNAs that couple to tracrRNA (gRNA, als0 called sgRNA). Individual gRNAs are bound by Cas protein effectors (e.g., Cas9). After a new phage infection with sequences matching a CRISPR spacer appears in the cell (lower right), specific Cas/gRNA complexes bind to viral DNA and cleave it. This Figure with the corresponding legend is reprinted from (Arroyo-Olarte, Bravo Rodríguez, &

Morales-Ríos, 2021). The permission to reuse this figure is under the terms of the Creative Commons CC BY license.

In 2012, scientists have tried to mimic this complex prokaryotic system and translate it into a simple genome editing tool as an RNA-programmable nuclease system by the fusion of the crRNA and the tracrRNA into a single synthetic gRNA which is comprised of a hairpin RNA that resembles tracrRNA linked to 20 bp sequence homologous to the target DNA (Jinek *et al.*, 2012). Also, the codon usage of Cas9 and RNaseIII genes of *Streptococcus thermophilus* were optimized and attached to nuclear localization signals to ensure nuclear localization in mammalian cells (Cong *et al.*, 2013; Mali *et al.*, 2013). In the CRISPR/Cas9 system derived from *Streptococcus thermophilus*, the target DNA must immediately precede a 5 prime NGG PAM (Jinek *et al.*, 2012), where Cas9 binds to the single guide RNA (sgRNA) scaffold and in the presence of the sgRNA complementary sequence and the PAM sequence Cas9 generates the DSB and this triggers the endogenous cellular DNA repair machinery (Jinek *et al.*, 2012) (**Figure 6**).

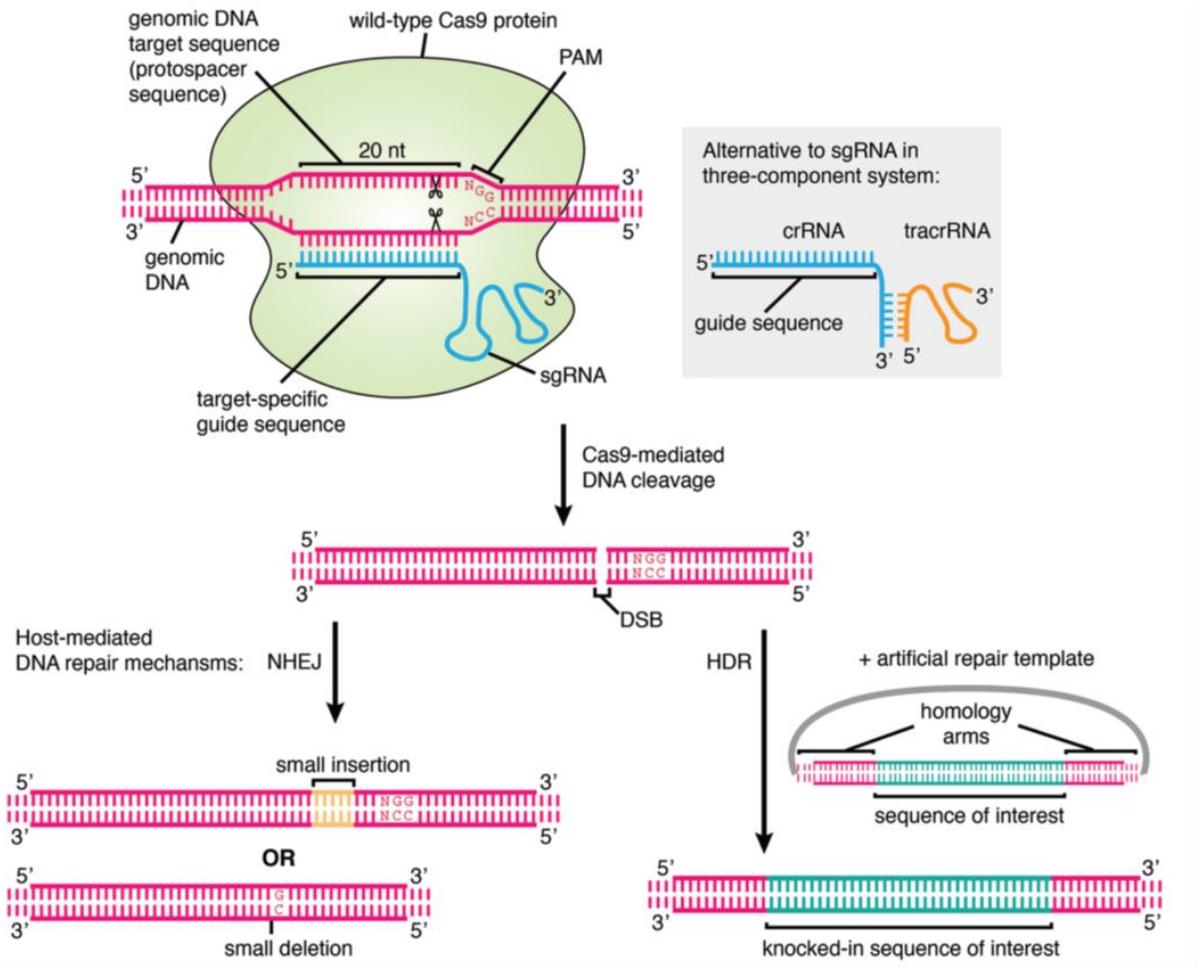


Figure 6: The mechanism of genome editing using CRISPR/Cas9.

The genomic DNA target, which must lie adjacent to a protospacer adjacent motif (PAM), is specified by a 20 nt user-generated guide sequence in the sgRNA or crRNA. The Streptococcus pyogenes PAM is shown. In the cell nucleus, Cas9 protein associates with the sgRNA or crRNA/tracrRNA and binds to the target sequence, cleaving both strands of the DNA at the site 3 nt upstream of the PAM. Cleavage results in a DSB which is repaired by host-mediated DNA repair mechanisms. In the absence of a repair template, error-prone NHEJ occurs which may lead to the formation of random short indels and thus frameshift mutations and disruption of gene function, and this represents the main method of CRISPR-mediated gene knock-out generation. If an artificial repair template is provided, for example on a plasmid containing a sequence of interest flanked by homology arms, then HDR may occur, leading to the integration of an exogenous DNA sequence at a specified genomic location. This is the basis for performing gene knock-in, tagging, and precise pre-specified insertions or deletions using CRISPR. If catalytically inactive Cas9 is used instead of wild-type Cas9, then the protein simply binds to the target locus and does not cleave the DNA, which can be used for transactivation or gene silencing. This figure with the correlated legend are re-printed from (Agrotis & Ketteler, 2015). The permission to reuse this figure is under the terms of the Creative Commons CC BY license.

CRISPR/Cas9 allowed cheaper, quicker, and easier design of human genome editing. Despite the simplicity and accessibility of CRISPR/Cas9 technology and its successful use for efficient

disruption of gene expression to simplify the creation of *in vitro* and *in vivo* disease models (Dow, 2015; Torres-Ruiz & Rodriguez-Perales, 2017), some aspects must be taken into account to improve the overall efficiency of this technology, like a delivery choice of this programmable nucleases and the possible side effects or cytotoxicity and economic accessibility of this technology (Jo, Suresh, Kim, & Ramakrishna, 2015). It has also to be taken into account the high rate of off-target effects which could be specific to different cell types (Fu *et al.*, 2013; C. Smith *et al.*, 2014), but several alternatives have been developed to overcome off-target effects and increase the sequence specificity at the target site through alteration of guide sequence by adding extra nucleotides at the 5 prime ends, and in addition to using nickases through mutant Cas9 to have one catalytic domain that cleaves only one strand (S. W. Cho *et al.*, 2014; Ran *et al.*, 2013).

1.3.1.3 Genomic safe harbors

Genomic safe harbors are intragenic or intergenic regions of the human genome that can incorporate transgenes to permit stable expression of this newly integrated DNA without any adverse effects on the host cell or organism (Sadelain, Papapetrou, & Bushman, 2011). Transgenesis of human cells is far used to study gene function in biomedical research and to treat disease through gene therapy. The first idea of genomic safe harbors for mammalian transgene insertion came in the late 1990s (Zambrowicz *et al.*, 1997) when Zambrowicz and colleagues found that one strain of mice known as ROSA beta geo26 expressed β galactosidase from a randomly inserted transgene and later they localized the transgene insertion site to chromosome 6. They found that the pups homozygous for the insertion developed normally and were fertile. Since that time *Rosa26* locus was recognized as a safe harbor and was used as a transgene insertion without any apparent harmful effects and permits stable gene expression.

Later, the genetic locus adeno-associated virus integration site1 (*AAVS1*) was described as a safe harbor for human cells (DeKelver *et al.*, 2010; Lombardo *et al.*, 2011). *AAVS1* locus is located on chromosome 19 (19q13.42) and was recognized as the special integration site of the human non-pathogenic adeno-associated virus (AAV) into the human gene *PPP1R12C*, which encodes phosphatase 1 regulatory subunit 12C protein and its function are not fully elucidated. It was noticed that the disruption of this gene by integration of AAV between exon 1 and intron 1 does not have any effect (Kotin, Linden, & Berns, 1992; Surks, Richards, & Mendelsohn, 2003). This locus was disrupted, and targeted insertions in several cell lines, including embryonic stem cells and induced pluripotent stem cells, maintain the expression of the inserted gene without any apparent harm to the examined cell lines (Hockemeyer *et al.*, 2008; J. R. Smith *et al.*, 2008).

AAVS1 was well characterized, and a specific TALEN pair was designed to induce DSB in the *AAVS1* locus, which is repaired by homologous recombination in the presence of a donor template, where the desired transgene sequences are placed between homologous recombination arms of the donor template. This approach can be used for the knock-in of the transgene. Recently this approach has become available commercially in an assay, where the homologous donor construct contains green fluorescent protein (GFP), and in addition, the transgene contains a marker for puromycin resistance. Treatment of transfected cells with puromycin will identify easily the clones that have successful integration of the transgene into *AAVS1*, and these puromycin-resistant cells will be green in fluorescence microscopy. In addition, the integration could be proven using primers spanning the five and three prime junctions by PCR analysis.

2 Research question, motivation and aim of the project

Using the power of next generation sequencing and with the great advances in the understanding of many genetic disorders, it was found that SBD constitute a clinically and genetically heterogeneous group of autosomal recessive diseases, caused by pathogenic changes in genes encoding enzymes of L-serine biosynthesis pathway *PHGDH*, *PSAT1*, and *PSPH*. However, due to the extreme rareness of these disorders, the number of reported cases and the mutational spectrum was very limited at the time when this project started, particularly regarding NLS. Moreover, previous work of the research group failed to identify a causative mutation in the aforementioned genes in a minority of cases with NLS, thus raising the possibility of further genetic heterogeneity.

SBD have been associated with a broad phenotypic spectrum ranging from non-specific developmental delay, neurodevelopmental impairment with microcephaly to the severe lethal disease known as NLS. With the discovery of the molecular basis of NLS, the question has been raised what the biological and pathophysiological basis for the wide clinical variability of SBD is? It has been supposed that the variability in the phenotype results from the degree of the residual enzyme activity, but no functional studies were available, espically for the mutants related with NLS.

The first aim of this study was to identify causative genetic variations in additional patients with a suspected SBD to expand the phenotypic and mutational spectrum of this disease group and to empirically establish correlations between genotype and phenotypic expression.

The second and major aim of the study was to establish an *in vitro* assay to validate the functional relevance of SBD-associated variants and to prove the assumption that the mutants' residual enzyme activities correlate with disease severity. Because PHGDH is the first enzyme

in serine biosynthesis pathway, this study focused on the modeling of PHGDH defects of both of NLS and SBDNL.

Using CRISPR/Cas9 as a genome editing tool, knock-out *PHGDH* cells were established and these cells have been used later to express *in vitro* transiently different mutants of *PHGDH*. Also, via gene targeting of the endogenous safe harbor locus *AAVS1* of these generated knock-out *PHGDH* cells, stable models of mutant *PHGDH* using TALEN system as another genome editing tool, were generated as knock-in *PHGDH*. This in turn, with the transiently transfected knock-out PHGDH cells were used to quantify the residual activity of mutant PHGDH and demonstrate its effect on the proliferation, viability of the cells.

The third aim was to identify possible novel genes for SBDs and to check the hypothesis of possible further genetic heterogeneity of NLS. To this end, exome or genome sequencing should be applied in patients with a convincing SBD (-like) phenotype who were tested negative for mutations in the known three genes.

3 Materials and Methods

3.1 Molecular diagnosis and genotype-phenotype correlations

3.1.1 Patients

Patients referred for molecular genetic testing for suspected NLS / SBD or with a diagnosis of such a disease established by whole exome sequencing (WES) were eligible for this study. All molecular testing was performed after obtaining informed consent according to the respective national regulations for genetic testing. Specific written parental permission was obtained for publication of photographs presented in this work.

The entire cohort included a total of 22 patients, from 19 unrelated families. Three patients that were first included for genetic screening, but who had no well-documented SBD phenotype, no images and only rudimentary clinical information, were excluded from further analysis and investigations after the molecular screening had revealed no mutation, these patients were considered rather not having an SBD.

The remaining study cohort included a total of 19 patients, 10 females and 9 males, from 15 unrelated families. Three of these families had a history of another probably affected offspring, but no details were available from these. The clinical data of all patients were collected using a questionnaire. Families were of various ethnic origin (**Table 1**). The majority of the patients were stillborn or had died shortly after the birth. In five cases, the pregnancy was terminated after severe malformations having been identified by fetal ultrasonography.

Table 1: Family information, auxology, and major malformations in the study cohort (15 families).

Patient / family number	1	2	3	4	5a	5b	6a	6b	7a	7b	8	9	10	11	12	13	14	15a	15b
Mutation detection method	TS	TS	TS	WES	TS	TS	WES	No gene analysis	WES	TS	WES	TS	TS	TS	TS	TS	TS	TS	TS
Gender	F	F	F	M	F	F	M	M	F	M	F	M	M	F	M	F	F	M	M
Parental consanguinity (specify)	Yes			1st degree	1st degree	1st degree	1st degree	1st degree			2nd degree	3rd degree		N/A	2nd degree				
Affected siblings				affected fetus in previous pregnancy	Sister	Sister	Brother	Brother	Male sibling fetus	Sister	Probably affected male sibling				Probably affected male sibling			Brother	Brother
Ethnic background	Caucasian (France)	Pakistani	N/A	Iranian	Iranian	Iranian	Turkish	Turkish	Caucasian (German)	Caucasian (German)	Arabian	Arabian (Algeria)	N/A	N/A	Turkish	N/A	Caucasian (Spain)	Tuvaluans	Tuvaluans
Gestational age at birth [weeks]	26+6	39+2	40	33	40	40	36	35	32+2	15	26	36+5	37+2	39	32	38+5	35	31+6	18+6
Birth weight (The number down represents the SD)	379 g (-2.5)	1988 g (-3.2)	2540 g (-2.0)	1179 g (-2.1)	1900 g (-3.7)	2200 g (-2.9)	2500 g (-1.7)	N/A	820 g (-2.4)	33 g	278 g (-3.2)	1330 (-3.8)	1371 g (-4.0)	2120 g (-2.9)	770 g (-2.6)	1507 g (-4.6)	N/A	1690 g (-0.2)	96.2 g
Body length at birth The number down represents the SD)	28 cm (-2.7)	42.5 cm (-3.1)	39 cm (-5.0)	31 cm (-4.9)	40 cm (-4.6)	39 cm (-5.0)	47cm (-1.1)	N/A	33 cm (-3.3)	9 cm	N/A	33 cm (-6.2)	38.6 cm (-4.8)	N/A	30 cm (-4.7)	N/A	N/A	43 cm (+0.5)	17.8 cm
OFC at birth The number down represents the SD)	18 cm (-4.5)	27.4 cm (-5.0)	28.2 cm (-4.7)	26 cm (-2.9)	30 cm (-3.4)	29 cm (-4.1)	30cm (-2.3)	N/A	24 cm (-3.4)	N/A	N/A	25 cm (-5.5)	24.6 cm (-6.1)	N/A	20.5 cm (-6.0)	26.4 cm (-5.4)	N/A	29 cm (-0.2)	11.6 cm
Survival /death	TOP	Liveborn, died on 1st d	Liveborn, died after 30 d	Stillborn	Liveborn, died after 27 d	Liveborn, died after 5 d	Alive age 4 y	Died at age 6.5 yr	Liveborn, died after 4 mo	TOP	TOP	Liveborn, died on 1st d	Liveborn, died after 5 d	Liveborn, died after 7 d	TOP	Liveborn, died on 1st d	Stillborn	Liveborn, died on 1st d	TOP
Cleft lip / cleft palate							CP	CP	CP							CP			CP
Central nervous system abnormalities	LIS, CH, CAL, HYD	LIS, CH, CAL, HYD	LIS, CH, CAL, HYD	LIS, CAL	N/A	N/A	CAL, HYD	N/A	CAL, HYD	N/A	LIS, CH, CAL, HYD	LIS, CAL, HYD	LIS, CH, CAL, HYD	N/A	LIS, CH, CAL	LIS, CH, HYD	N/A	CH	CH
Limb and skeletal anomalies	CON	CON, EDE	CON, EDE	CON, EDE	CON	CON	CON, EDE	CON	CON	EDE	CON, EDE	CON, EDE	CON	CON, EDE	CON	CON, EDE	CON, EDE	CON, EDE	CON
Skin abnormalities	ICH	RDERM	mild ICH	RDERM	ICH	ICH	mild ICH	N/A		N/A		ICH	RDERM	RDERM	ICH	RDERM	RDERM	RDERM	ICH

Genitourinary abnormalities	GEN		KID	GEN			GEN	N/A	GEN			GEN, KID	GEN		KID		GEN	GEN,, KID	KID
Additional anomalies	Cataract	Cataract			Cataract	Cataract	Cataract	Cataract	CDA, PD	PEX GM		VSD, sacral agenesis	Cataract		Cataract , AVSD, HRHS, situs inversus	MH			

Estimation of gestational age in patient 15a was based on fetal size at the first scan, which correlated with a gestational age of 21+4 weeks; it is therefore probably underestimated.
Atrioventricular septal defect (AVSD); callosal hypoplasia or agenesis (CAL); Congenital duodenal atresia, (CDA); cerebellar hypoplasia (CH); cleft lip and cleft palate (CLP); joint contractures (CON); cleft palate (CP); day (d); edema / swelling of hands and/or feet (EDE); female (F), gastrointestinal anomalies (GA); genital hypoplasia / anomalies (GEN);gut malformation (GM); grams (g); hypoplastic right ventricle (HRHS); hydrocephalus / enlarged ventricles (HYD); ichthyosis (ICH); kidney anomalies (KID) , lissencephaly spectrum (LIS); male (M); month (mo); myocardial hypertrophy (MH), N/A not available; occipitofrontal circumference (OFC); pectus excavatum (PEX) polydactyly (PD); restrictive dermopathy (RDERM); standard deviation (SD); termination of pregnancy (TOP); targeted sequencing (TS); ventricular septal defect (VSD); whole exome sequencing (WES); year (yr).

3.1.2 Molecular genetic analysis

DNA was extracted from the blood samples of the patients using Chemagic Magnetic Separation Module in combination with Chemagic DNA Blood Kit or using QIAcube robotic workstation in combination with the QIAamp DNA Blood Mini Kit according to the manufacturer's protocols. For families 4 and 11, where only formalin-fixed, paraffin-embedded tissue (FFPE) from the affected fetus was available, DNA was extracted using QIAamp®DNA FFPE Tissue Kit according to the manufacturer's protocol. From three families, DNA samples of affected fetuses (patients 5a / 5b, 12, and 14) were not available, and therefore parental DNA was tested assuming heterozygous carrier status. In family 5, a chorionic villus sampling DNA (CVS DNA) was analyzed to provide a prenatal diagnosis in this family with a previous child clinically diagnosed with NLS.

In most samples, targeted Sanger sequencing was performed for all coding exons and flanking intronic regions of the known causative genes for NLS [*PHGDH* (ENST00000369409), *PSAT1* (ENST00000376588), *and PSPH* (ENST00000395471)]. Sequence data were generated using the BigDye Terminator v3.1 Cycle Sequencing Kit and an automated capillary sequencer (ABI 3500). The Sequence Pilot software (JSI medical systems, Germany) was used for analysis. Due to the very poor quality of the DNA extracted from the FFPE tissue of patients 4 and 11, no full mutation screening of the three genes of interest was possible, but only amplification of a very short fragment of exon 7 of *PHGDH* and exon 4 of *PSAT1* encompassing the most common NLS-associated variant in this gene was successful, respectively. In patient 13, owing to a very low volume and concentration of the available DNA sample, amplification of the genomic DNA sample was done using the GenomiPhi V2 DNA Amplification Kit, according to the manufacturer's instructions. In four cases (families 4, 6, 7, and 8) the molecular diagnosis was made by diagnostic (trio) exome sequencing in the labs of our collaborators (Abdelfattah

et al., 2020). All variants identified in the three genes were analyzed using the online prediction tools MutationTaster (http://www.mutationtaster.org/), Meta-SNP (http://snps.biofold.org/meta-snp/**)**, and CADD (https://cadd.gs.washington.edu/). The human splicing finder (http://www.umd.be/HSF3/) was used to predict the effect of the splice site variants. Mutated protein stability was also analyzed *in silico* using an integrated predictor for protein stability change upon a single mutation (http://predictor.nchu.edu.tw/istable/**)**.

All the variants were submitted to the genetic Leiden Open Variation Database (LOVD) (https://databases.lovd.nl/).

3.1.3 Structural analysis and molecular modeling

The structural analysis of the protein variants was done by Professor Heinrich Sticht from the Institute of Biochemistry, Friedrich-Alexander-Universität Erlangen, Germany. This analysis was performed based on the crystal structures of PHGDH (PDB code: 2G76) and PSAT1 (PDB code: 3E77). Missense changes were modeled with SwissModel (Guex & Peitsch, 1997) and RasMol (Sayle & Milner-White, 1995) was used for structure analysis and visualization.

3.2 Functional analysis of *PHGDH* mutants

Introductory comment regarding the naming of the different generated cell types and the mutated enzymes in this study

The acronym **wtHEK** has been used in this study to refer to the wild type of HEK cells, whereas the abbreviation **koHEK** has been used to refer to the generated knock-out *PHGDH* HEK293 cells which were generated by CRISPR/Cas9.

The letter **T** or **S** in a lowercase has been added to the abbreviation koHEK to indicate that these generated koHEK cells were transiently or stably transfected with different constructs of *PHGDH*, respectively.

Whereas the first letter, number, and the second letter added after S or T indicate the location and names of the altered amino acids in the resulted PHGDH produced by transiently or stably transfected koHEK cells.

The table below **Table 2** clarifies the names of all the generated cell lines and its corresponding enzymes. In addition, one should mention that all the produced enzymes had the same name as the cells that produced it on the diagrams and in the text of this thesis for ease of clarification.

Table 2: The names of all the generated cell lines in this study.

used names of the generated cell lines	**variant of *PHGDH***	**the corresponding expressed PHGDH**
koHEK-tWT	WT	overexpressed WT
koHEK-tR54C	c.160C>T	p.Arg54Cys
koHEK-tR163Q	c.488G>A	p.Arg163Gln
koHEK-tT213M	c.638C>T	p.Thr213Met
koHEK-tV261M	c.781G>A	p.Val261Met

koHEK-tE265*	c.793G>T	p.Glu265*
koHEK-tE265K	c.793G>A	p.Glu265Lys
koHEK-tE265E	c.795G>A	p.Glu265Glu
koHEK-tG377S	c.1129G>A	p.Gly377Ser
koHEK-tV425M	c.1273G>A	p.Val425Met
koHEK-tQ433*	c.1297C>T	p.Gln433*
koHEK-tV490M	c.1468G>A	p.Val490Met
koHEK-sWT	WT	overexpressed WT
koHEK-sT213M	c.638C>T	p.Thr213Met
koHEK-sV261M	c.781G>A	p.Val261Met
koHEK-sE265*	c.793G>T	p.Glu265*
koHEK-sE265K	c.793G>A	p.Glu265Lys
koHEK-sE265E	c.795G>A	p.Glu265Glu

In this study, the stable transfected cells' name can be appended with PGK or with CMV to refer to the used cloning donor vector (pAAVS1D-**PGK**-MCS-EF1a-copGFP-T2A-Puro or pAAVS1D-**CMV**-MCS-EF1a-copGFP-T2A-Puro); respectively to get the stable transfected *PHGDH* koHEK cell lines (knock-in *PHGDH* HEK cell lines).

For the nomenclature of the mutation, one letter code without p. is used in the following functional part of this study (e.g., V490M instead of p.Val490Met).

3.2.1 Vectors

All the vectors used for the experiments are listed in **Table 3** and all the vector maps are shown in **Figure 7**, **Figure 8**, **Figure 9**, and **Figure 10**.

Table 3: Names of the used vectors in the functional study.

Name of the vector	Catalog number	Supplier	Comment
GeneArt®CRISPR Nuclease vector	A21174	Thermo Fisher Scientific Inc.	Provided with GeneArt® CRISPR Nuclease OFP Reporter Kit.
TOPO-TA cloning vector	450641	Thermo Fisher Scientific Inc.	Provided with TOPO TA Cloning® Kit.
pTagGFP-C vector	FP191	Evrogen JSC	----------
pZT-AAVS1-L1/R1 TALE-Nuclease vectors	GE601A-1	System Biosciences LLC	The Kit AAVS1 TALE Nuclease (GE600A-1) conntains AAVS1 HR Donor vector with PGK promoter and two AAVS1 TALE-Nuclease vectors, which designed to efficiently cut the *AAVS1* locus. Also it contains a positive control donor vector.
AAVS1 HR cloning Donor vector (pAAVS1D-PGK-MCS-EF1a-copGFP-T2A-Puro)	GE602A-1	System Biosciences LLC	PGK promoter was replaced with a strong CMV promoter by Dr. Sönke Weinert to have two cloning Donor vectors with different promoters.
pAAVS1D-CMV-RFP-EF1a-copGFP-T2A-Puro	GE603A-1	System Biosciences LLC	This positive control donor vector and the cloning donor vectors contain homologous sequences to ~800bp upstream and downstream of the AAVS1 TALE-Nuclease targeting sites.

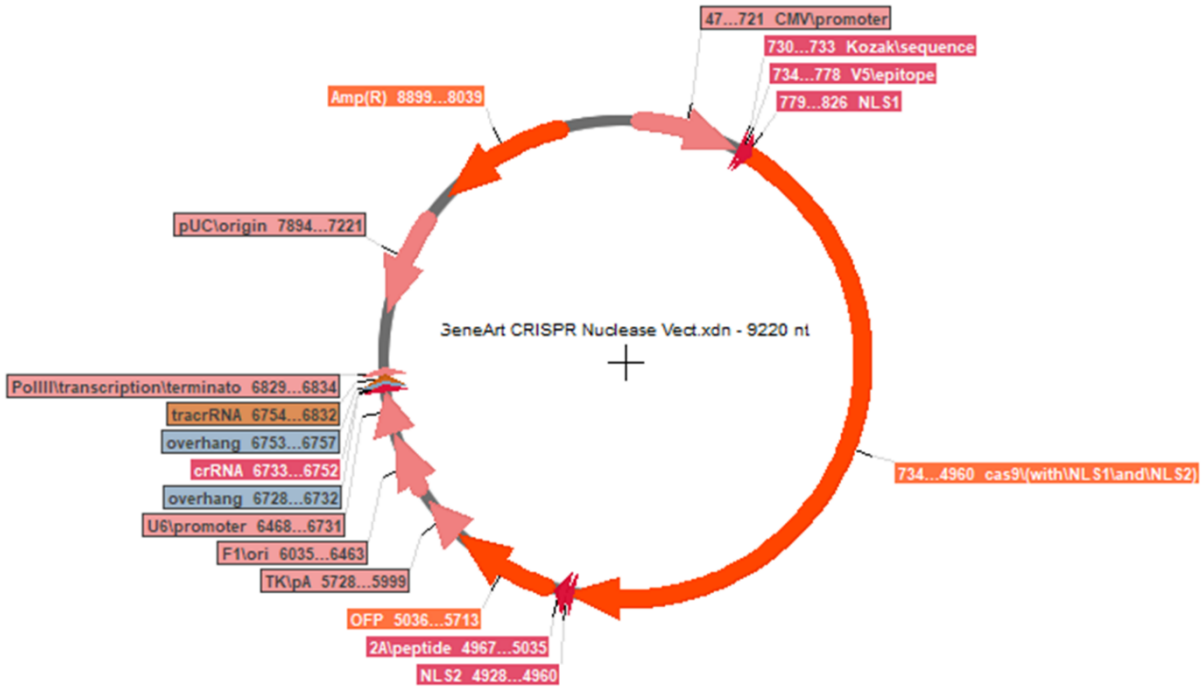


Figure 7: GeneArt®CRISPR nuclease vector map.

The vector is pre-linearized with 5 base pair overhangs for easy cloning of my designed double-stranded DNA oligo that encodes crRNA. This vector offers all in one expression vector consisting of both a Cas9 nuclease expression cassette and gRNA cloning cassette. Cas9 is fused to orange fluorescent protein (OFP) via a self-cleaving 2A peptide linkage allowing the expression of both from the same mRNA. Cas9 is directed to the nucleus by nuclear signals (NLS1 and NLS2). The map of the vector was constructed using Serial Cloner (http://serialbasics.free.fr/Serial_Cloner.html).

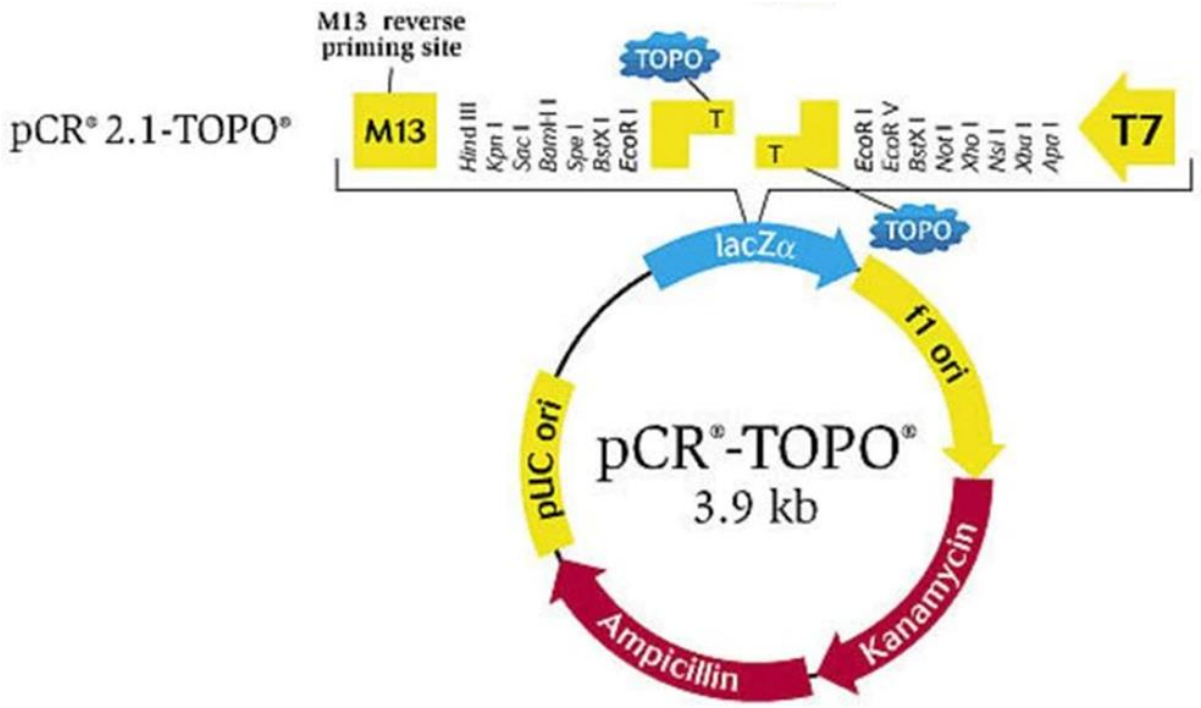


Figure 8: Topo-TA cloning vector.

This vector provides a highly efficient, 5-minute, one-step cloning strategy for the direct insertion of Taq polymerase-amplified PCR products into a plasmid vector. No ligase or PCR primers containing specific sequences are required. This vector is linearized with single 3 prime thymidine (T) overhang for TA cloning and with Topoisomerase I bound to the vector. It is known that Taq polymerase adds a single deoxyadenosine (A) to the 3 prime ends of PCR products and that allows PCR inserts to ligate efficiently with the vector without ligase. The figure is reprinted from the protocol of the TOPO-TA Cloning® Kit (Thermo Fisher Scientific Inc.).

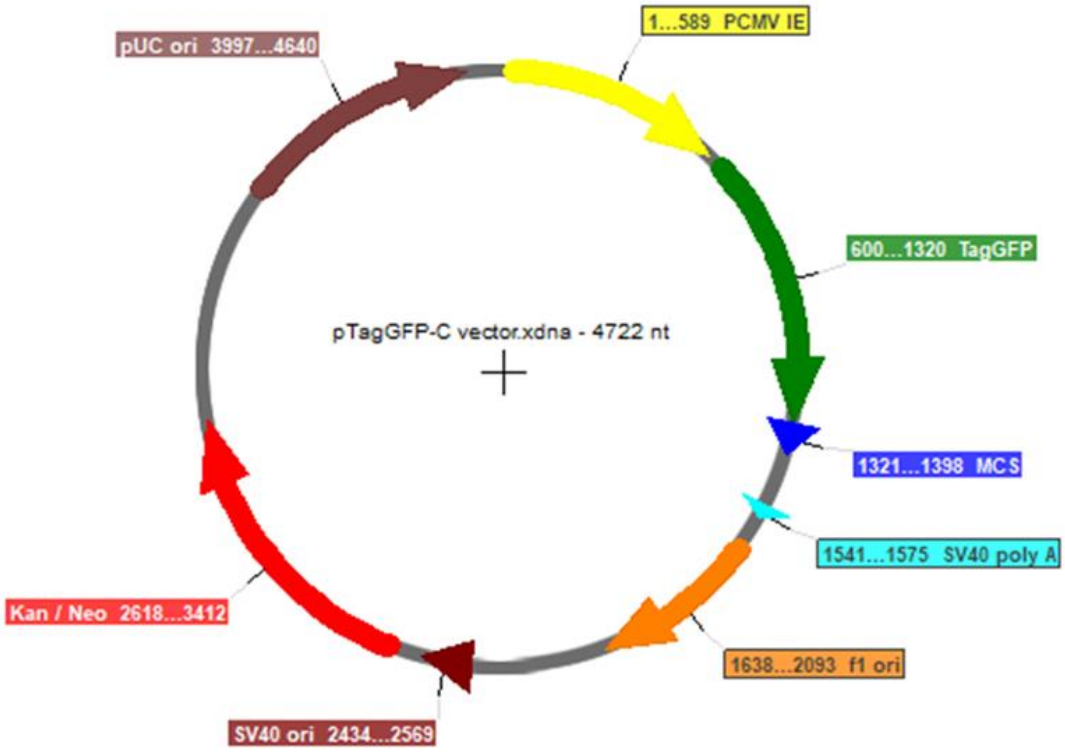


Figure 9: pTagGFP-C expression vector.

This vector encoding green fluorescent protein TagGFP and allows generation of fusions to the TagGFP C-terminus and expression of TagGFP fusions in mammalian cells. The map of the vector was constructed using Serial Cloner (http://serialbasics.free.fr/Serial_Cloner.html).

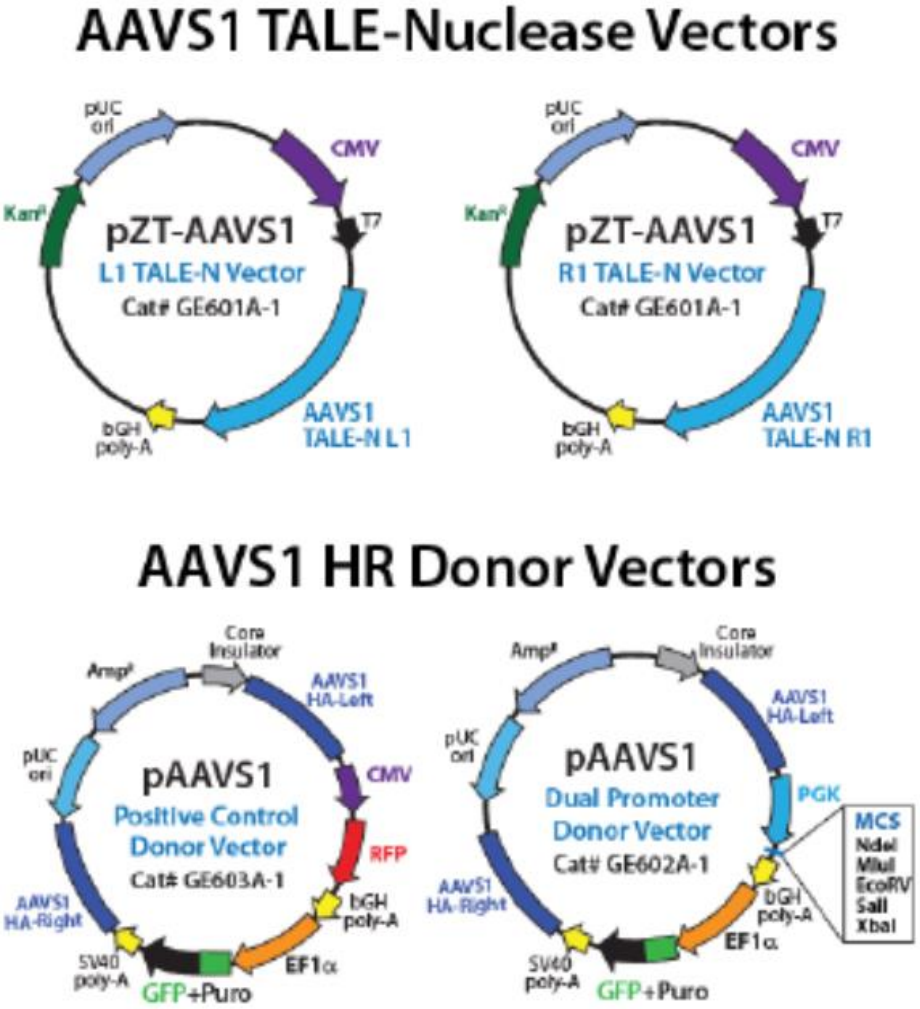


Figure 10: Vector maps for all the vectors provided in AAVS1 TALE-nuclease Kit.

pZT-AAVS1-L1/R1 vectors cut efficiently at a defined site on the AAVS1 locus, whereas pAAVS1D-CMV-RFP-EF1a-copGFP-T2A-Puro positive control donor vector was used to validate TALE-Nuclease cutting and subsequent HDR in target koHEK by detection of RFP&GFP markers and puromycin resistance, since it contains homologous sequence to ~ 800 bp upstream and downstream of the AAVS1 TALE-Nuclease targeting sites. pAAVS1D-PGK-MCS-EF1a-cop GFP-T2A-Puro donor cloning vector contains homologous sequence to ~ 800 bp upstream and downstream of the AAVS1 TALE-Nuclease targeting sites and HDR was detected by GFP marker and puromycin resistance. The promoter PGK in this donor cloning vector was replaced with cytomegalovirus (CMV) promoter to use two donor vectors with two different promoters in this study. The figure is reprinted from the protocol of AAVS1 TALE-Nuclease Kit (System Bioscience LLC).

3.2.2 Lab equipments

All the used lab types of equipment are listed in **Appendix A Table 1**.

3.2.3 Consumables

All the consumables are listed in **Appendix B Table 1**.

3.2.4 Reagents and Kits

All the used reagents, kits, and buffers are listed in **Appendix C Table 1**, **Appendix C Table 2**, and **Appendix C Table 3**.

3.2.5 Oligonucleotides

All the used oligonucleotides are listed in **Appendix D Table 1**, **Appendix D Table 2**, **Appendix D Table 3**, **Appendix D Table 4**, **Appendix D Table 5**, **Appendix D Table 6**, **Appendix D Table 7**, and **Appendix D Table 8**. The corresponding oligonucleotides for amplification, sequencing, and target-specific sequence in the genome were purchased from Thermo Fisher Scientific Inc. (Darmstadt, Germany). Oligonucleotides for junction polymerase chain reaction (PCR) were purchased from Eurofins Genomics Germany GmbH (Ebersberg, Germany). All the used oligonucleotides for cloning and qPCR were from Integrated DNA Technologies, Inc. (Coralville, USA). Oligonucleotides were designed with the help of Primer3 version 4.0.0 (https://bioinfo.ut.ee/primer3-0.4.0/) and were analyzed using IDT's OligoAnalyzer™ Tool (https://eu.idtdna.com/calc/analyzer). 100 µM (=100 pmol/µl) was the concentration of all stock solutions, but 2.5 µM of each primer was used in the working solution for PCR.

3.2.6 Antibodies

All the used antibodies are listed in **Table 4**.

Table 4: Antibodies used in this study.

Antibody (Cat. Number)	Specificity	Company
Anti-Cyclin D1 (92G2)	Rabbit monoclonal antibody	Cell Signaling Technology Inc.; Massachusetts, USA
Anti-Vinculin (E1E9V)	Rabbit monoclonal antibody	Cell Signaling Technology; Massachusetts Inc., USA
Anti-PHGDH (PA5-54360)	Rabbit polyclonal antibody	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Peroxidase-AffiniPure goat anti-rabbit IgG (H+L)	This secondary antibody reacts with whole molecule rabbit IgG	Biozol Diagnostica Vertrieb GmbH, Eching, Germany

3.2.7 Cells (eu- and prokaryotes)

Human embryonic Kidney 293 wild type cells (wtHEK) were purchased from DSMZ-German Collection of Microorganisms and Cell Cultures GmbH (Braunschweig, Germany) and were used as the eukaryotic cell line to perform genome editing as the basis for the functional part of this study.

For cloning experiments, two different *E. coli* cell lines were used as listed in **Table 5**.

Table 5: Name and genotype of the used prokaryotes cells in the study.

Name	Supplier	Genotype
One Shot® TOP10 Competent Cells	Thermo Fisher Scientific Inc.; Darmstadt, Germany	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(araleu)7697 galU galK rpsL (StrR) endA1 nupG
Stellar™ Competent Cells	Takara Bio Europe SAS; Saint-Germain-en-Laye, France	F–, endA1, supE44, thi-1, recA1, relA1, gyrA96, phoA, Φ80d lacZΔ M15, Δ (lacZYA - argF) U169, Δ (mrr - hsdRMS - mcrBC), ΔmcrA, λ–

3.2.8 Workflow of the functional study

The following diagram **Figure 11** summarizes the workflow regarding the functional study of *PHGDH*.

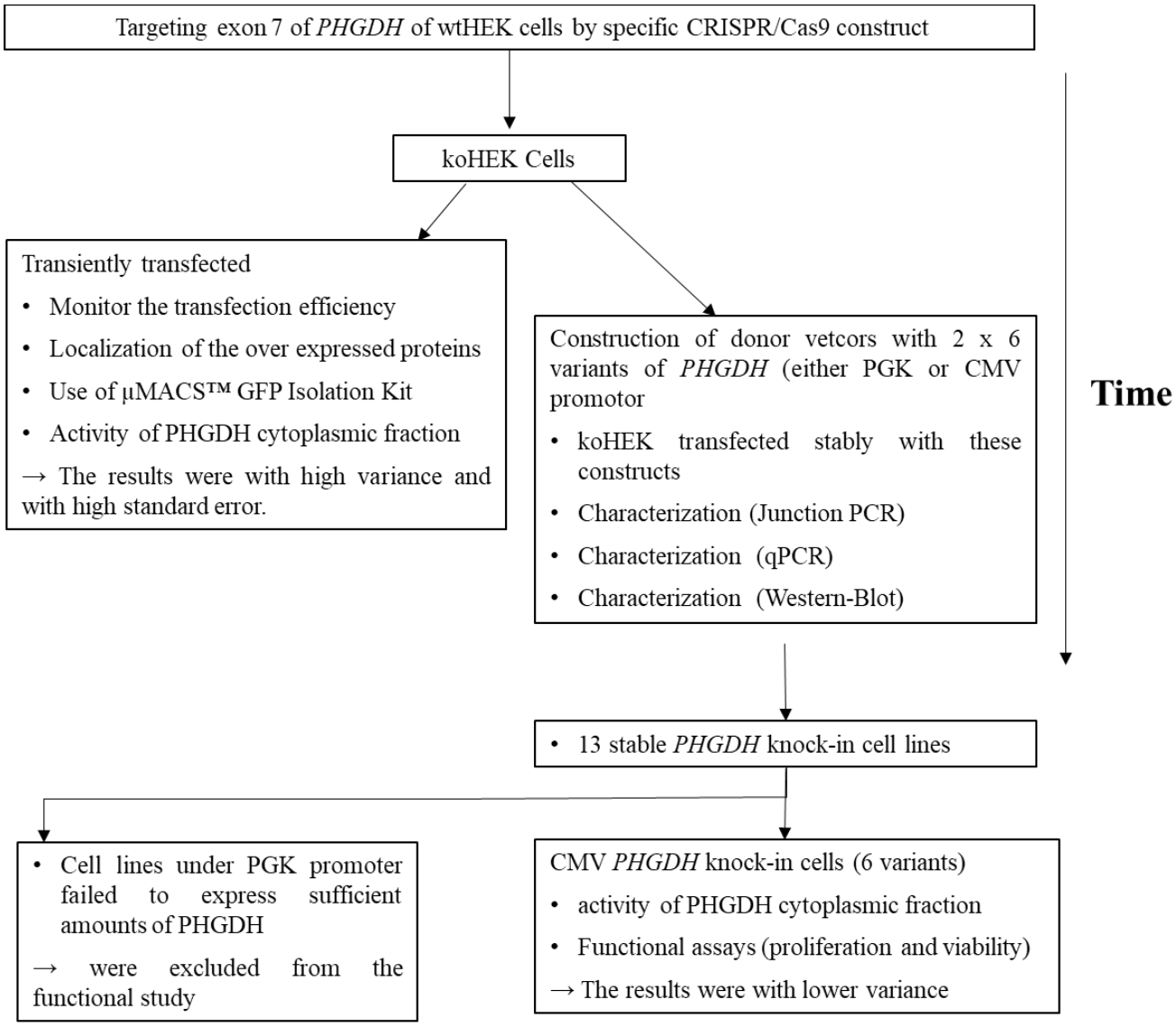


Figure 11: Diagram summarizes the workflow regarding the functional study of PHGDH.

Targeting exon 7 of PHGDH of wtHEK cells by specific CRISPR/Cas9 construct generated PHGDH-knock-out cells (koHEK cells). These koHEK cells were transiently transfected with an empty vector of pTagGFP-C and with 12 different constructs of pTagGFP-C, which expressed 11 different mutants and WT of PHGDH. The transfection efficiency and the localization of the expressed proteins were monitored by fluorescence microscopy. µMACS™ GFP Isolation Kit was used to purify the recombinant proteins to use it to measure the enzyme activity, but the activity was lost during the purification, then cytoplasmic fraction lysis was used to estimate the residual activity of PHGDH for all these transgenic cells lines→ The results were with high variance and with high standard error.

koHEK cells were transfected stably by 6 different constructs of donor vectors either with CMV or PGK promoter and with the donor vector which was provided by the kit. and 13 different knock-in PHGDH HEK cell lines were generated (6 under CMV promoter expressing wt and 5 different mutations of PHGDH, 6 under PGK promoter expressing wt and the same 5 mutations of PHGDH and one cell line transfected with the control donor vector). The integration of the donor vectors was validated by junction PCR. The expression of the inserted PHGDH with 6 different variants was monitored by western blot and for some variants by qPCR. Cell lines under the PGK promoter failed to express sufficient amounts of PHGDH and for this reason,were excluded from the functional study. Knock-in PHGDH under CMV promoter cell lines were used to estimate the residual activity of PHGDH. These knock-in PHGDH HEK cell lines were used also to conduct cell biology assays like proliferation and viability of the cells.

3.2.9 Generation of knock-out *PHGDH* HEK293 cells

To generate a *PHGDH*-deficient cell line, wtHEK cells were transfected with *PHGDH* targeting CRISPR nuclease vector. To this end, two single-stranded DNA oligonucleotides were designed, one encoding crRNA sequence to target a specific site of exon 7 of *PHGDH* and the other on the reverse strand. The online tools CHOPCHOP (https://chopchop.cbu.uib.no/**)** and CRISPRdirect (https://crispr.dbcls.jp/**)** were used for the design (Labun, Montague, Gagnon, Thyme, & Valen, 2016; Naito, Hino, Bono, & Ui-Tei, 2015). Sequences of the oligonucleotides used in this study are provided in **Appendix D Table 1**. The GeneArt®CRISPR nuclease vector was used according to the manufacturer's protocol to obtain a *PHGDH* targeting CRISPR nuclease vector, which was purified from the transformed One Shot® TOP10 competent cells using GeneJET Plasmid Maxiprep Kit according to the manufacturer's protocol. wtHEK cells were transfected with the *PHGDH* targeting vector using Lipofectamine® 3000 reagent according to the manufacturer's protocol and as described later in paragraph 3.2.12.1. Single cells were sorted by serial dilution ending with the final concentration of 2 cells in 1 ml and 100 µl was distributed in each well of 96 well culture plates. After initial cell cultivation, genomic DNA (gDNA) was extracted using QIAcube robotic workstation in combination with the QIAamp DNA Blood Mini Kit according to the manufacturer's protocols. The targeted region in exon 7 of *PHGDH* was amplified by PCR using Platinum® Taq DNA Polymerase and the amplicons were cloned with the use of TOPO-TA cloning® Kit to obtain single allelic clones

for DNA sequencing. About eight single allelic TOPO-TA clones were sequenced per individual cell clone. Cell clones with bi-allelic FS mutations predicting a premature stop codon and thereby functional knock-out of *PHGDH* (koHEK cells) were subcultivated and cryopreserved at -150 °C for later characterization and usage. A schematic illustration of the strategy for genome editing and the allele-specific sequence is presented in **Appendix E Figure 1**. The generated koHEK cell line was used in this work for transient and stable transfection experiments with selected *PHGDH* mutations.

3.2.10 RNA isolation and cDNA synthesis

Total RNA was extracted from healthy human lymphoblastoid cell lines using the RNeasy® Mini Kit according to the manufacturer's instructions. Quality and quantity of RNA were estimated using RNA ScreenTape Analysis and RNA samples were stored at -80 °C. Synthesis of first-strand complementary DNA (cDNA) from purified RNA was performed with the SuperScript III Reverse Transcriptase Kit, with the addition of Oligo(dT) 12-18 Primer and [d(N)6] random hexamers. All steps were performed on ice and only RNase-free ddH_2O was used. The quality of the resulting cDNA was evaluated using *GAPDH* control primers (**Appendix D Table 2**) as a positive control. cDNAs were stored at -20 °C and were later used to amplify the open reading frame (ORF) of *PHGDH* using primers with a 5' overhang including recognition sites for cloning into the vectors (**Appendix D Table 3**).

In the subsequent steps of the experimental procedure, RNA from the wtHEK, koHEK, and the generated stable transgene cell lines (as described later in paragraph 3.2.12.3) were extracted using InviTrap®Spin Cell RNA Mini Kit according to the manufacturer's protocol. To prevent the degradation of RNA 1 U/µl RiboLock RNase inhibitor was added to the isolated RNA. The RNA concentration was determined using NanoDrop 2000/2000c UV-Vis spectrophotometer

and to remove any remaining DNA, samples were subjected to a one-hour DNaseI treatment at 37 °C using RQ1 RNase-free DNase. Later DNaseI was inactivated by the addition of 1 µl of DNase-Stop solution and heat inactivation at 70 °C for 10 min. 1 µg of the resulting RNA was reverse transcribed into cDNA using a High Capacity cDNA Reverse Transcription Kit using random hexamers primers according to the manufacturer's instructions. For the subsequent step, the cDNA samples were diluted by TE-buffer (1X) pH 8.0 to get 25 ng/µl to use for quantitative real-time PCR (qPCR).

3.2.10.1 Polymerase chain reaction (PCR)

PCR was used several times in this work for amplification all coding exons of the three genes (*PHGDH*, *PSAT1*, *PSPH*) for the purpose of sequencing or to amplify the ORF of *PHGDH* by using primers containing restriction recognition sites for subsequent cloning into vectors. PCR was used also to insert specific mutations to the ORF of *PHGDH* by the amplification using primers containing the desired mutants.

To this end, different templates, gDNA, cDNA and vectors were subjected to PCR. As it is shown in **Appendix A Table 1**, several devices were used, and several programs were optimized to get the desired amplicons for the subsequent steps. Taq DNA Polymerase or Pwo DNA Polymerase were used according to the manufacturer's protocols supplemented with primers, template, Mg^{2+} and buffer to achieve the desired amplification. For DNA samples of low concentration, Platinum® Taq DNA Polymerase or Platinum SuperFi DNA Polymerase were used. The reactions were either performed in 96 well PCR plates or in 8 well strips.

3.2.11 Molecular cloning and molecular biology methods

3.2.11.1 Restriction cloning

Primers with a 5' overhang including a recognition site for KpnI in the forward primer and a recognition site for BamHI in reverse primer were used to amplify the full-length *PHGDH* ORF from cDNA of a control human lymphoblastoid cell line by polymerase chain reactions (sequence of primers in **Appendix D Table 3**). Amplicons were separated via gel electrophoresis (agarose 0.8%) and bands corresponding to the correct PCR product were excised and purified by QIAquick® Gel Extraction Kit according to the manufacturer's protocol. Amplicons were subjected to restriction digest with KpnI and BamHI and were then ligated into the multiple cloning site of the expression vector pTagGFP-C (FP191, Evrogen JSC) in frame with the TagGFP. The vector was prior digested and purified with the same restriction enzymes. Ligation was performed using T4 DNA Ligase at 22 °C for one hour and afterward, the reaction was inactivated at 65 °C for 10 min and was kept on ice or at -20 °C. Stellear™ competent *E. coli* cells were thawed for 30 min on ice and about 7 µl of the ligation products were added. The solution was incubated on ice for 30 min and then subjected to heat shock for 30 sec at 42 °C. The solution was immediately incubated on ice for 2 min. Subsequently, 250 µl of warm SOC medium was added and the solution was incubated for one hour with shaking at 230 rpm at 37 °C. The components of the SOC medium are shown in **Appendix C Table 3**. The bacterial suspension was plated and cultivated on kanamycin (30 µg/ml) LB-Agar plates for 24 h. Isolated colonies were picked and cultured in a lysogeny broth medium containing kanamycin for 24 h and the constructs were isolated via QIAprep®Spin Miniprep Kit according to the manufacturer's protocol. ZymoPURE™II Plasmid Maxiprep Kit was used to enable high yield and concentration of both empty and *PHGDH* wt constructs. To remove any remaining

ions from plasmid preparations, ethanol precipitation was done overnight at -30 °C, using 3 M sodium acetate solution (pH 5.2) and cold 96% (*v/v*) ethanol. On the next day, the pellet was washed several times with 70% ethanol (*v/v*) and was left to dry with an open lid. Then DNA pellet was resuspended in sterilized water. The concentration and purity of the vectors were determined spectrophotometrically via the absorption at 260 nm and 280 nm and the constructs were stored at -20 °C until usage.

The resulting construct pTagGFP-C-*PHGDH* wt was used later as a template to introduce 11 different mutants to the ORF of *PHGDH* using these-mutants-containing primers by either classical four primer mutagenesis PCR or QuickChange Lightning Multi Site-Directed Mutagenesis Kit according to manufacturer's protocol. Following the same strategy, which was elucidated above, 12 mutated constructs of the pTagGFP-C-*PHGDH* vector were obtained, ready to use in the subsequent steps of this work. The sequence of the oligonucleotide mutated primers with the name of the variants is shown in **Appendix D Table 3**.

3.2.11.2 Gibson cloning

Phosphoglycerate kinase PGK promoter in the pAAVS1D-PGK-MCS-EF1a-copGFP-T2A-Puro cloning donor vector provided from AAVS1 TALE-Nuclease Kit was replaced by a CMV promoter provided by Dr. Sönke Weinert (Experimental Cardiology, Faculty of Medicine, OVGU Magdeburg) to get additional cloning donor vector pAAVS1D-CMV-MCS-EF1a-copGFP-T2A-Puro with a strong promoter. Both cloning donor vectors were linearized by restriction enzyme SalI according to the manufacturer's protocol. Specific primers for *PHGDH* with 15 bp extensions (5') that are complementary to the ends of the linearized vectors and contained the start codon in the forward primer and the stop codon in the reverse primer, were designed to amplify just the ORF of *PHGDH* from the construct pTagGFP-C-*PHGDH* wt,

which was used as a template (primer sequences provided in **Appendix D Table 4**). The integrity and the size of the amplicons were verified on agarose gel (0.8%). Cloning reactions were done using In-Fusion®HD Cloning Kit, were incubated for 15 min at 50 °C then placed immediately on ice. 2 µl of the reactions were subsequently used to transform Stellar™ competent cells. After an initial preculture, the cells were plated on LB-Agar plates supplemented with ampicillin (100 µg/ml). Colonies were picked and the same strategy as mentioned above was followed to get a high yield and concentration of both cloning donor *PHGDH* wt constructs. The same strategy was followed using five mutated pTagGFP-C-*PHGDH* constructs as a template (c.638C>T, c.781G>A, c.793G>T, c.793G>A, c.795G>A) to obtain ten mutated pAAVS1 TALE donor constructs, each of them with PGK promoter (pAAVS1D-PGK-*PHGDH*-EF1a-copGFP-T2A-Puro) and with CMV promoter (pAAVS1D-CMV-*PHGDH*-EF1a-copGFP-T2A-Puro).

3.2.11.3 Validation of all the constructs

All the generated constructs, encoding either *PHGDH* wt gene or mutants were verified by diagnostic restriction digest of 100-300 ng of the purified constructed vectors with the restriction enzymes for one hour at 37 °C, followed by a run of the digestion reaction on agarose gel 1 % in parallel with uncut vector to distinguish the band resulting from the backbone from that resulting from the insert (wt or mutated ORF of *PHGDH*). This analytical diagnostic provides indirect information about the sequence. Subsequently, Sanger sequencing of all the constructs was done to validate the sequence of the insert of either *PHGDH* wt or mutants to exclude the errors, which might have been introduced by PCR. (The sequence of the designed primers to span exon-exon boundaries is shown in **Appendix D Table 5**). Bacterial glycerol stock of all the validated constructs was prepared and frozen at -80 °C in a cryovial for long-

term storage through mix a determined amount of fresh bacterial overnight culture with the same amount of 50% glycerol solution.

3.2.11.4 Genomic DNA extraction

gDNA from the generated knock-out and knock-in *PHGDH* HEK cell lines were extracted using the Invisorb® Spin Blood Mini Kit from the cell suspension in PBS according to the manufacturer's protocol.

3.2.11.5 Junction PCR

Junction PCR was used in this work to validate the integration of donor constructs either of pAAVS1D-PGK-*PHGDH*-EF1a-copGFP-T2A-Puro or pAAVS1D-CMV-*PHGDH*-EF1a-copGFP-T2A-Puro specifically at the *AAVS1* locus. For this purpose, specific primer pairs, that flank the 5-prime *AAVS1* and 3-prime *AAVS1* homology arms right and left of the insert, were designed (primer sequences are shown in **Appendix D Table 6**). The junction PCR reaction was performed by Platinum™ II Hot-Start Green PCR Master Mix using the touchdown PCR protocol as was recommended in the manufacturer's protocol. Amplicons were analyzed on a 1 % agarose gel in 1X TAE buffer to confirm the integration of the homology arms of the donor vectors via the appearance of the expected size of the amplicon.

3.2.11.6 Real Time PCR

qPCR was used in this work to confirm and compare the expression of the inserted external (*PHGDH*) in the *AAVS1* locus and the internal expression of *PHGDH* in each of wtHEK, koHEK and the generated stable cell lines either with CMV or PGK promoter cell lines. qPCR was done using SYBR® Premix Ex Taq™ II (Tli RNaseH Plus). In this assay, DNA-binding dye (SYBR Green II) was used to monitor the amplification of the *PHGDH* sequence. This dye

binds nonspecifically to double-stranded DNA (dsDNA) and exhibits little fluorescence when it is free in the solution, but its fluorescence increases up to 1000-fold when it binds dsDNA. Therefore, the overall fluorescent signal from a reaction is proportional to the amount of dsDNA present and will increase as the target is amplified. To compare the internal *PHGDH* sequence from the *PHGDH* construct, the forward primer was designed to locate in exon 11 of *PHGDH* and the reverse primer in the 3-prime untranslated region of *PHGDH*. This primer pair amplified only the internal *PHGDH*, whereas using the same forward primer with another reverse primer located in the AAVS1 HR cloning donor constructs, specifically amplified the external *PHGDH* provided by the vectors (primer sequences provided in **Appendix D Table 7**). For each pair of primers, an RNA and non-template control (water control) were used instead of the cDNA to exclude any DNA contamination. The comparison between external and internal *PHGDH* expression was done relative to the wtHEK and normalized to the housekeeping genes (ß-2-microglobulin and *HPRT1*) in the same sample in the exponential phase of the qPCR. To evaluate assay performance, the standard curve was constructed by running a serial dilution of cDNA and plotting the log of the dilution factor against the threshold cycle (C_T) value obtained during the amplification of each dilution. Then the equation of the linear regression line, along with the coefficient of determination (R^2), were used to evaluate the optimization of my qPCR assay. The R^2 value of a standard curve represents how linear the data are and gives a measure of the variability across assay replicates since the qPCR was carried out in duplicate for each sample and each primer pair. Amplification efficiency (E) was also calculated from the slope of the standard curve using the following formula: $E = 10^{-1/slope}$. This value was presented as a percentage and was close to 100% as the best indicator of a robust, reproducible assay. The analysis was performed using Bio-Rad CFX-Manager™ Software.

3.2.12 Cell culture and Transfection

wtHEK, koHEK, and the generated stable transfected HEK cells were cultured in 100 x 20 mm sterile dishes containing Dulbecco's Modified Eagle Medium (DMEM) with high glucose (4.5 g/l glucose, L-glutamine, and 25 mM HEPES without pyruvate) plus 10% (v/v) fetal bovine serum (FBS) and 1% (v/v) antibiotic/antifungal mix and were incubated at 37 °C, 5% CO_2, 95% humidity. All the reagents are provided in **Appendix C Table 1**.

3.2.12.1 *PHGDH* targeting of wtHEK cells by *PHGDH* targeting vector

To transfect wtHEK cells with the constructed CRISPR nuclease vector, Lipofectamine®3000 Reagent was used. Approximately 200,000 cells were cultured in 6 well plates to be 70-80% confluent at the time of the transfection. The reagent Lipofectamine®3000 was diluted in Opti-MEM medium. Similarly, 5 µg of constructed DNA *PHGDH* targeting CRISPR nuclease vector was diluted in the same medium and P3000 reagent was added. After that, diluted DNA was added to the diluted Lipofectamine®3000 reagent and was incubated for 5 min at room temperature, and then this DNA-lipid complex was added to wtHEK to transfect them. If gRNA and nuclease are delivered, a DSB can be induced and a mutagenic NHEJ event can occur. Transfected cells can be visualized by the expression of OFP. Cells were analyzed on a flow cytometer 72 h post-transfection to estimate the transfection efficiency and the data were evaluated using the FlowJo software (BD Life Sciences, New Jersey, USA). To estimate the CRISPR/ nuclease-mediated cleavage efficiency, the GeneArt®Genomic Cleavage Detection Kit in combination with the Tapestation system D1000 ScreenTape assay was used according to the manufacturer's protocols.

Single-cell clones were sorted in 96 well culture plates, and then were subcultivated and passaged to 48, 24, 12, 6 wells, and 10 cm plates. After validation of their sequence, as described

in paragraph 3.2.9, suitable cell clones were cryopreserved to be used in the subsequent steps as functional koHEK cells.

3.2.12.2 Transfection of koHEK cells with pTagGFP-C constructs for transient expression

Two million koHEK cells were plated 48 h before transfection on 10 cm plates. 5 µg for each expression vector (empty vector, wt *PHGDH* vector, and 11 mutated *PHGDH* vectors) as illustrated in **Appendix D Table 3** were used to transfect the cells using standard calcium phosphate precipitation. The components of the transfection solution are provided in **Appendix C Table 3**.

On the day of the transfection, the confluency was adjusted to 70-90%, the culture medium was aspirated and replaced by DMEM containing chloroquine 25 µM half hour before the transfection, and the cells were further cultivated in the incubator. The DNA was mixed directly with ddH_2O and a concentrated solution of calcium chloride 2.5 M (final concentration 250 mM), which was then added dropwise to a phosphate buffer (2X HBS buffer pH 7.05) to form a fine precipitate. Aeration of the phosphate buffer while adding the DNA-$CaCl_2$ solution helps to ensure that the formed precipitate is fine since clumped DNA will not adhere to or enter the cell efficiently. On the two following days, the success of the transfection was monitored with a fluorescence microscopy. PolyFect® transfection reagent was used one time according to the manufacturer's protocol to achieve better imaging conditions (three-dimensional imaging) using the fluorescence microscopy.

3.2.12.3 *AAVS1* targeting of koHEK cells and generation of stable knock-in *PHGDH* HEK cell lines

The same strategy as described in paragraph 3.2.12.2 was used to co-transfect koHEK cells with 3 µg of each of TALEN Nuclease pZT-AAVS1-L1/R1 pairs, which were designed to efficiently cut *AAVS1* locus at a defined site, in combination with 10 µg of each of the twelve constructed donor vectors (as mentioned in paragraph 3.2.11.2). All these donor vectors contained a homologous sequence to ~800 bp upstream and downstream of the *AAVS1* TALEN targeting sites, which were flanking the ORF of *PHGDH* either wt or mutated. Co-transfection of the TALEN pairs and the donor vectors leads to TALEN cutting of the *AAVS1* locus and subsequent HDR provided by the donor vectors. This approach produced cell lines with twelve different constantly expressed PHGDH mutants and wt (knock-in PHGDH HEK cell lines). Because the PHGDH of the donor verctors were incorporating a GFP and puromycin resistance as selection markers. These markers were further used to select positive knock-in cells by puromycin selection and to monitor the success of the selection by fluorescence microscopy. Twenty-four h post-transfection, the transfection media was removed and the cells were split 1:3 in complete growth media containing puromycin (2 µg/ml). The selection was continued until only GFP-positive cells were in the culture. One of three plates was saved for characterization of TALEN modified and HDR recombinant cells by junction PCR and the second for RNA isolation and then subsequent qPCR. After the complete selection of the generated cell lines by puromycin, the third plate was cryopreserved to be used in the subsequent experiments.

3.2.12.4 Analysis of PHGDH activity and biological function

3.2.12.4.1 Proliferation assay (BrdU assay)

A colorimetric BrdU immunoassay, which measures pyrimidine analog 5- Bromo-2'-deoxyuridine (BrdU) incorporation instead of thymidine into the DNA of proliferating cells, was used to test the proliferation rates of the different cell lines, wtHEK, koHEK, and six knock-in *PHGDH* cell lines. Only cells with the transgene under the CMV promoter were used in this experiment because their expression was closer to the natural PHGDH expression. These eight different cell lines were plated into 96 well culture plates (11,000 cells/well) and grown in serine-deficient medium (MEM) at 37 °C for 48 h. Each cell line was analyzed with 8 technical replicates in 8 wells. After 48 h of culture, BrdU was added to the cells, and cultures were reincubated again for 2 h at 37 °C, 5% CO_2, and 95% humidity. To quantify cell proliferation activity, BrdU incorporation was measured by immunoassay using the spectrophotometry Synergy™ HT Microplate reader, with absorbance at 450 nm, following the manufacturer‘s instructions. The experiment was repeated three times.

3.2.12.4.2 Viability assay

CellTiter96®AQ_{ueous} Non-Radioactive Cell Proliferation Assay was used to assess metabolic activity and viability of the same cell lines, which were described above in 3.2.12.4.1. To this end, we used the same amounts of cells as described above plated into 96 well culture plates with 8 replicates each, and cultured them in MEM for 48 h. Subsequently, a combined reagent from the assay was added to the cells and the cultures were reincubated again for 3 h. This reagent is composed of a novel tetrazolium compound (MTS) and an electron coupling reagent phenazine methosulfate. MTS is bioreduced by dehydrogenase enzymes in metabolically active cells into an aqueous, soluble formazan product in a culture medium. For quantification, the

absorbance of the formazan was measured at 490 nm using a Synergy™ HT Microplate reader. The experiment was repeated four times.

3.2.12.4.3 Enzyme activity assay

PHGDH Activity Assay Kit was used to monitor the activity of PHGDH mutants expressed transiently and stably in koHEK cells (13 different constructs of pTagGFP-C-*PHGDH* with transient expression and six different cell lines with stable expression under the CMV promoter). Empty vector and overexpressed wt served as controls. wtHEK and koHEK cells were also included. In all cases, four million cells were lysed with the provided lysis buffer and centrifuged at 17,000 g for 10 min to remove cell debris. Two and a half volumes of 4.32 M ammonium sulfate solution were added to 100 µl of the supernatant and the mixture was incubated on ice for 1 h, followed by centrifugation at 17,000 g at 4 °C for 10 min. The resulting pellet was resuspended again in 100 µl of assay buffer, then 20 µl samples were pipetted into a 96-well clear plate followed the final volume per well was adjusted to 50 µl using assay buffer and subsequently 50 µl of the reaction mixture was added. Three wells for each variant with reaction mixture and one well for the same variant with background mixture were included. The plate was measured immediately at 450 nm in kinetic mode measuring every 60 s for the next 60 min at 37 °C. Two-time points in the linear range of the curve were chosen to calculate the PHGDH activity as recommended in the manufacturer's protocol. The specific activity of PHGDH was calculated using the same equation relative to the content of total protein in each well. Relative activity was then determined by normalization of the results to the specific activity measured for overexpressed wt PHGDH. This assay was repeated four times for each of the transiently expressed proteins and the stably expressed proteins.

Trials, in the beginning, were also performed to isolate recombinant proteins using µMACS™ GFP isolation Kit through their tagged GFP proteins. This µMACS™ technology uses anti-GFP microbeads for the magnetic labeling and direct isolation of tagged proteins from different protein sources. The µMACS microbeads in this kit are a colloidal suspension of extremely small (50 nm in diameter) super-paramagnetic particles, and they can bind specifically to GFP allowing fast and effective binding to GFP-proteins. Magnetically labeled proteins are retained on a µ column placed in the magnetic field of a µMACS separator. To remove non-specific interacting molecules, stringent washing steps were applied. Then to get native PHGDH protein, GFP-proteins were eluted by pH shift using triethylamine, pH 11.8. However, owing to the change in the pH of the resulting native recombinant proteins, it was difficult to use these proteins for enzyme assay since they lost their activity during the purification and this assay was never tried use later with the purified proteins.

3.2.13 Protein extraction and Western blotting

Cells were lysed in the prepared lysis buffer supplemented with a protease inhibitor as shown in **Appendix C Table 3**, followed by centrifugation to get proteins. Quantitation of protein was performed using the Quick Start™ Bradford Protein Assay Kit. Equal amounts of proteins were loaded and separated on 7.5% SDS-polyacrylamide gels. Two extra 12% SDS-polyacrylamide gels were used to confirm the absence of any truncated form of PHGDH in knock-out stably transfected with a nonsense mutation (koHEK-sE265*).

PHGDH protein was detected using a rabbit polyclonal anti-PHGDH, whereas vinculin, which was used as a loading control, was detected using monoclonal rabbit anti-vinculin. All the antibodies were prepared in 3% bovine serum albumin. Peroxidase-AffiniPure goat anti-rabbit IgG was used as a secondary antibody. The chemiluminescence reagent Western Lightning®

Plus-ECL was used for visualization. Gel analysis and quantification were done using ImageJ (Schneider, Rasband, & Eliceiri, 2012) and the normalization was done always to koHEK cell line transfected transiently or stably with wt PHGDH (koHEK-tWT or koHEK-sWT, respectively) to calculate the relative PHGDH protein expression.

3.2.14 Statistical analysis

All measurements were performed in at least three biological replicates. If repeated measurements of the same sample were used (technical replicates), the result for the samples was calculated as the average of the measurements. Statistical analyses were performed using GraphPad InStat (GraphPad Software Inc., San Diego, USA). One-way Analysis of Variance (ANOVA) and Dunnett multiple comparisions test was done. The level of significance was set at $p \leq 0.05$. Numbers of biological replicates for each experimental condition are provided in the respective figure legends.

4 Results

4.1 Mutation screening

In 15 families with NLS, variants in either *PSAT1* or *PHGDH* were identified that were considered as likely causative (**Table 6**, **Appendix F Table 1**). In one of these families (family 4) with abnormal fetuses occurring in two branches of the family, which was previously reported negative for a mutation in the three genes for NLS (family 11 in Acuna-Hidalgo *et al*. (Acuna-Hidalgo *et al.*, 2014)), exome sequencing of additional family members revealed a likely disease-causing *PHGDH* variant in one branch of the family (**Figure 12**).

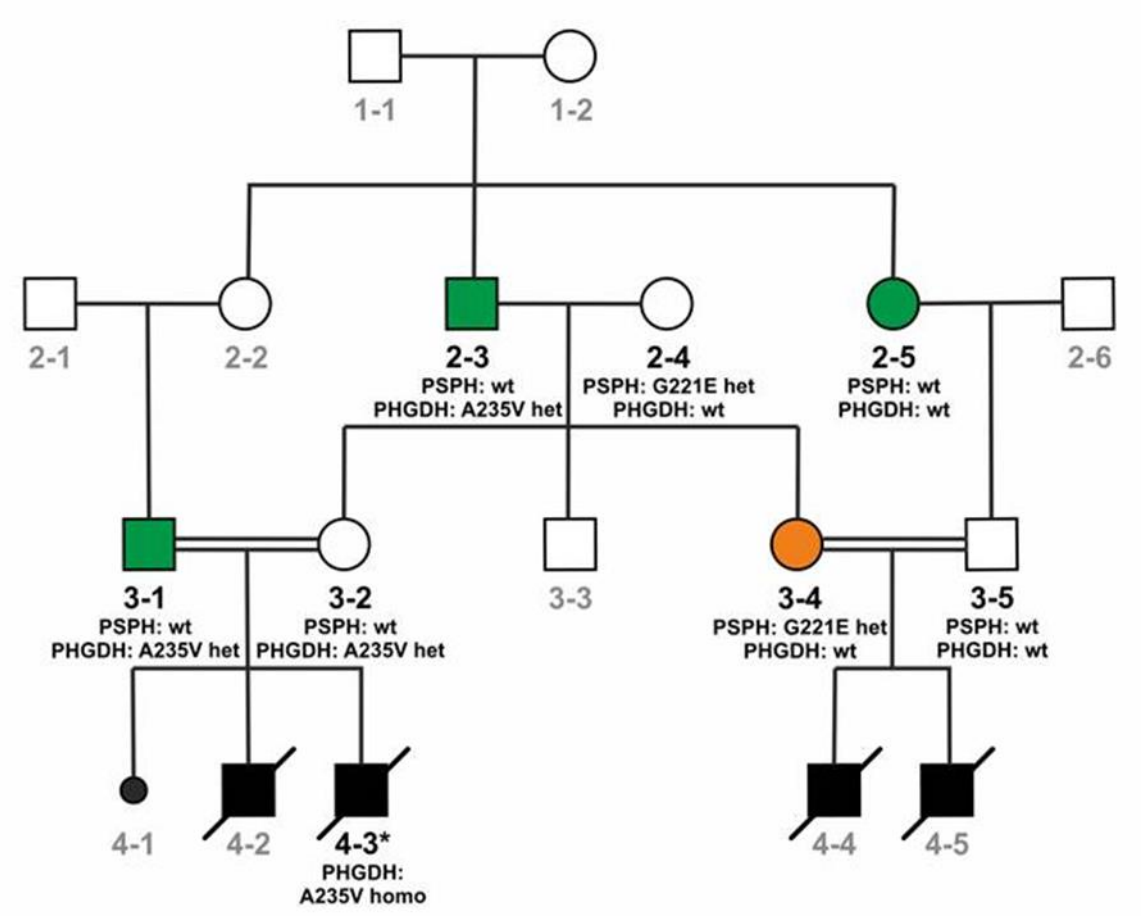


Figure 12: The pedigree of family 4.

This family previously published as family 11 by Acuna-Hidalgo et al. (Acuna-Hidalgo et al., 2014). A multi-consanguineous family was initially considered as affected with NLS in two branches of the family. All affected individuals were stillborn or neonatal deaths. Materials for genetic testing were available from individuals whose ID numbers are printed in black, while no material was available from the others (ID numbers in grey). From individual 4-3 only FFPE tissue material was available (*) allowing the extraction of very poor-quality DNA. Initial whole exome sequencing (WES) was performed on individual 3-4 (orange circle) who was considered as an obligate carrier for the presumed familial mutation responsible for the disease. The heterozygous PSPH variant, c.662G>A (p.Gly221Glu, G221E), which was identified in this person, was inherited from her mother (2-4) but not present in any of the other family members who were expected to be carriers (assuming autosomal recessive inheritance). The cause of the disease in the family remained unsolved and the possibility of further genetic heterogeneity of NLS was discussed (Acuna-Hidalgo et al., 2014). In the second approach performed as part of this work, WES was performed on three additional family members considered obligate carriers (green boxes and circles). This analysis revealed a heterozygous variant in the PHGDH gene in individual 2-3 and 3-1, while 2-5 was wild type (wt) for this variant. Additional targeted Sanger sequencing also confirmed the carrier status for this variant in 3-2 and homozygosity in the affected fetus 4-3, while the variant was not present in the other branch of the family (right). The disease in the affected fetuses of 3-4 and 3-5 was retrospectively defined as not compatible with NLS.

In total, the disease was attributable to *PSAT1* in 11 unrelated families and to *PHGDH* in 4 (**Appendix F Table 1**, **Figure 13**). *PSPH* variants were not observed in this cohort. A total of 13 different presumably disease-causing variants were identified in those two genes. Nine of them were novel (previously not reported in NLS). Two *PSAT1* variants were recurrent in this cohort: one of the previously reported variants (c.296C>T, p.Ala99Val) was observed in four unrelated families and one novel splice acceptor change (c.870-1G>T) was observed in three unrelated families.

The novel *PHGDH* and *PSAT1* variants were either absent from the gnomAD database (https://gnomad.broadinstitute.org/**)** or present at a very low frequency, compatible with the expected carrier frequency for a very rare autosomal recessive disease. The highest allele frequency of 1.52e-4 was recorded for the recurrent *PSAT1* variant (c.296C>T, p.Ala99Val). Three of the novel missense variants received a formal classification of a variant of uncertain significance (VUS) according to American College of Medical Genetics (ACMG) recommendations (Richards *et al.,* 2015), while all other observed variants were classified as either pathogenic or likely pathogenic (**Table 6**). The results of the analysis of all the novel variants by various *in silico* prediction tools are summarized in **Table 6**. Affected individuals

from six families with known parental consanguinity were homozygous for the variant considered as causative, while three families had different compound heterozygous constellations (families 7, 10, and 13). In the non-consanguineous family 3, a previous SNP microarray that was done in an external lab showed loss of heterozygosity for the entire length of chromosome 1, indicating that homozygosity for the *PHGDH* variant resulted from uniparental disomy (UPD) of chromosome 1. The parental origin of the *PHGDH* mutation could not be further assessed, because parental samples were not available. In one family (family 11), where paternal DNA was unavailable and the fetal material derived from FFPE tissue insufficient for a full screening, only the maternally inherited *PSAT1* variant could be identified, allowing the attribution of this case to PSAT1 deficiency (**Appendix F Table 1**), although the genotype was not fully elucidated.

The distribution of novel and previously described disease-associated variants is shown in **Figure 13**.

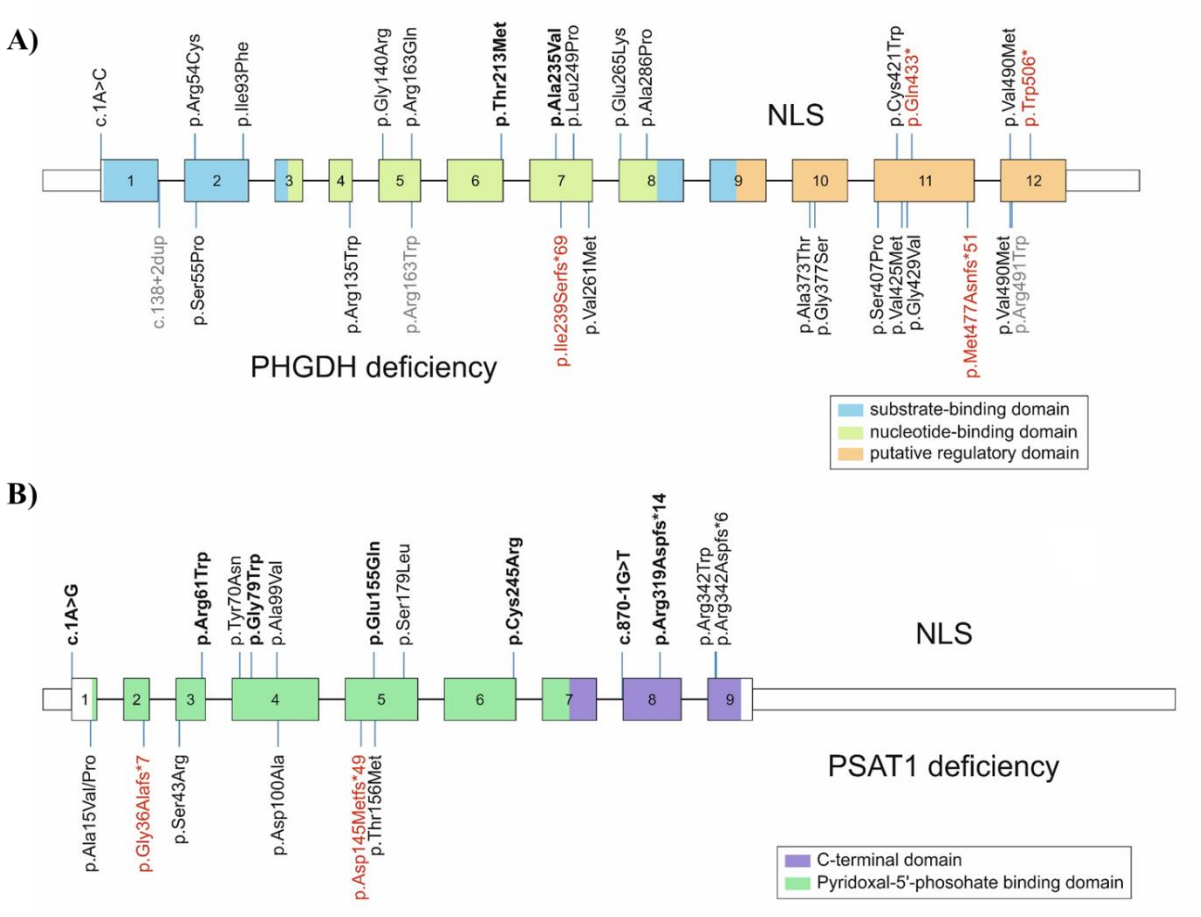


Figure 13: Distribution of mutations along the PHGDH and PSAT1 genes.

*Schematic representation of variants in the genes PHGDH (**A**) and PSAT1 (**B**). Exons are to scale (larger rectangles represent coding regions, lower represent untranslated regions (UTR)) joined by a continuous line (introns, not to scale). Different colors refer to the different domains of the respective proteins as indicated. Disease–associated variants are shown with their locations along the genes. Variants observed in NLS are depicted above the diagrams, while variants observed in the nonlethal forms of PHGDH and PSAT1 deficiency, respectively, are shown below the diagrams. Novel variants are depicted in bold. Alterations predicting a loss-of-function (putative null alleles) are printed in red color. Variants whose functional impact could not be classified are printed in grey color. Prediction of different domains of PHGDH are based on the knowledge of the 94% similarity between rat and human 3-phosphoglycerate dehydrogenase (PHGDH) and the paralogous enzyme of Escherichia coli, sharing 30% identical amino acids with human 3-PGDH (Achouri et al., 1997; H. M. Cho et al., 2000; Klomp et al., 2000; Pind et al., 2002; Tabatabaie et al., 2009). Prediction of different domains of PSAT1 are based on the knowledge of the extensive homology between E.coli SerC and phosphoserine aminotransferase of rabbit and human that share 93,5% of their amino acid sequence (Hester et al., 1999; van der Zel, Lam, & Winkler, 1989). NLS, Neu-Laxova syndrome, PHGDH, phosphoglycerate dehydrogenase, PSAT1 phosphoserine aminotransferase, UTR, untranslated region.*

Notably, most NLS-associated *PHGDH* variations predicting missense changes affect the nucleotide-binding domain and the substrate-binding domain of the protein, respectively, whereas the majority of the *PHGDH* variants previously observed in non-lethal PHGDH deficiency are located in the C-terminal regulatory domain (**Figure 13A**). In contrast, pathogenic *PSAT1* variants associated with non-lethal PSAT1 deficiency are located in the first 5 exons of the gene, but the number of alleles is too small to draw a definite conclusion and no obvious phenotype-specific distribution was observed (**Figure 13B**).

Table 6: PHGDH and PSAT1 variants observed in the present cohort.

Gene	Variant	Consequence on RNA / protein	gnomAD#	Meta-SNP	Mutation Taster	CADD‡	iStable	3D modeling	ACMG	LOVD DB-ID
PHGDH	c.160C>T	p.Arg54Cys	5/0	Disease causing	Disease causing	25.3	Decrease (0.6923)	Near substrate binding site, may cause reduced substrate affinity	LPATH (PM2, PM3, PP2, PP3)	PHGDH_000021
PHGDH	c.488G>A	p.Arg163Gln	Not found	Disease causing	Disease causing	33.0	Decrease (0.8418)	Predicted to affect dimer stability	LPATH (PS4, PM2, PP2, PP3)	PHGDH_000002
PHGDH	**c.638C>T**	**p.Thr213Met**	2/0	Disease causing	Disease causing	28.8	Decrease (0.7492)	Predicted to hamper NAD+ binding	VUS (PM2, PP2, PP3)	PHGDH_000018
PHGDH	**c.704C>T**	**p.Ala235Val**	Not found	Disease causing	Disease causing	27.4	Decrease (0.8443)	Predicted to hamper NAD+ binding	VUS (PM2, PP2, PP3)	PHGDH_000020
PSAT1	**c.1A>G**	**p.Met1?**	Not found	N/A	N/A	24.8	N/A	N/A	LPATH (PVS1, PM2)	PSAT1_000009
PSAT1	c.129T>G	p.Ser43Arg	Not found	Disease causing	Disease causing	23.0	Decrease (0.7772)	Predicted to affect dimer stability	LPATH (PS4, PM2, PP2, PP3)	PSAT1_000017
PSAT1	**c.181C>T**	**p.Arg61Trp**	24/0	Disease causing	Disease causing	29.9	Decrease (0.8558)	Predicted to destabilize the enzyme fold by steric clash	LPATH (PS3, PM2, PP1, PP2, PP3)	PSAT1_000010
PSAT1	**c.235G>T**	**p.Gly79Trp**	Not found	Disease causing	Disease causing	31.0	Increase (0.6263)	Predicted to cause steric	VUS (PM2, PP2, PP3)	PSAT1_000011

								clashes with PLP		
PSAT1	c.296C>T	p.Ala99Val	43/0	Disease causing	Disease causing	24.2	Decrease (0.7616)	Predicted to lead to protein instability	PATH (PS3, PS4; PM2, PM3, PP1, PP2, PP3)	PSAT1_000016
PSAT1	**c.463G>C**	**p.Glu155Gln**	1/0	Disease causing	Disease causing	28.4	Decrease (0.8194)	Predicted to affect PLP binding	LPATH (PM2, PM3, PP2, PP3)	PSAT1_000014
PSAT1	**c.733T>C**	**p.Cys245Arg**	8/0	Disease causing	Disease causing	33.0	Increase (0.5808)	Predicted to cause steric clashes with V39 and M42, may reduce dimer stability	LPATH (PS3, PM2, PP2, PP3)	PSAT1_000013
PSAT1	**c.870-1G>T**	**Splicing, p.?**	5/0	N/A	N/A	26.5	N/A	N/A	PATH (PVS1, PM2, PM3)	PSAT1_000012
PSAT1	**c.955delA**	**p.Arg319Aspfs*14**	Not found	N/A	Disease causing	35.0	Decrease (0.7891)	N/A	PATH (PVS1, PM2, PP1)	PSAT1_000015

Variants refer to the reference sequences of PHGDH (NM_006623.3) and PSAT1 (NM_058179.3). Novel mutations are printed in bold.
 # Representation in gnomAD is given as number of observed alleles / numbers of homozyotes.
*‡ CADD PHRED score: CADD (http://cadd.gs.washington.edu/) v1.4 PHRED-like (-10*log10(rank/total)) scaled C-score: ranking a variant relative to all possible substitutions of the human genome (8.6x10^9). A scaled C-score ≥10: variant belongs to 10% most deleterious variants; C-score ≥20: variant belongs to 1% most deleterious variants*.

4.2 Structural analysis and molecular modeling

To better understand the effects of the novel variants on the protein structure and functional consequences, structure analysis and molecular modeling of the novel NLS-associated missense changes were conducted in comparison to known PHGDH and PSAT1 deficiency-associated mutants. For both, PHGDH and PSAT1, the functionally active enzyme proteins are dimers, and their dimeric configuration is known to be essential for their proper function (John, 1995; Mishra, Ali, Nozaki, & Bhakuni, 2010).

Taken together, the structural analyses suggest that the majority of the previously reported and newly detected *PHGDH* and *PSAT1* missense variants cluster at distinct sites of the protein structure. The major region that is affected by disease-associated variants is the substrate/cofactor binding site of the respective proteins, suggesting that the respective mutants directly affect enzymatic activity. A second hot spot region in both proteins is the subunit interface in the homodimeric enzymes, indicating that the respective variants indirectly affect activity via a reduced dimer stability. However, it must be noted that some additional variants were also located at other sites outside of those clusters. These changes may affect enzymatic activity by different mechanisms, like for example a reduced overall stability of the domain structure (**Table 6**, and supporting results related to each of **Appendix E Figure 2**, **Appendix E Figure 3** and **Appendix E Figure 4**).

4.3 Phenotype analysis and genotype-phenotype correlations

The physical findings in 19 affected individuals with *PHGDH* and *PSAT1* variations are summarized in **Table 6** and are provided in more detail in **Appendix F Table 1**. External phenotypic features of selected patients are shown in **Figure 14**.

All cases except the two affected children from family 6 represented pre- or perinatal deaths with an unambiguous clinical diagnosis of NLS. Four individuals died shortly after the birth, two were stillborn, and in five cases, the pregnancy was terminated upon the detection of severe fetal anomalies by ultrasound or after genetic confirmation of recurrence of the disease in the fetus (patients 1, 7b, 8, 12; and 15b). The remaining individuals diagnosed as typical NLS survived up to maximum age of 4 months. The median age at death of the liveborn patients with NLS (excluding family 6) was 5 days. In contrast, the affected children from family 6 exhibited a less severe phenotype which was interpreted as intermediate between mild NLS and a very severe expression of non-lethal PSAT1 deficiency. One child died at age 6 years and the other one (patient 6a) was still alive at age 4 years. All affected individuals showed significant intrauterine growth restriction (-1.7 SD to -4.6 SD for term or near-term newborns).

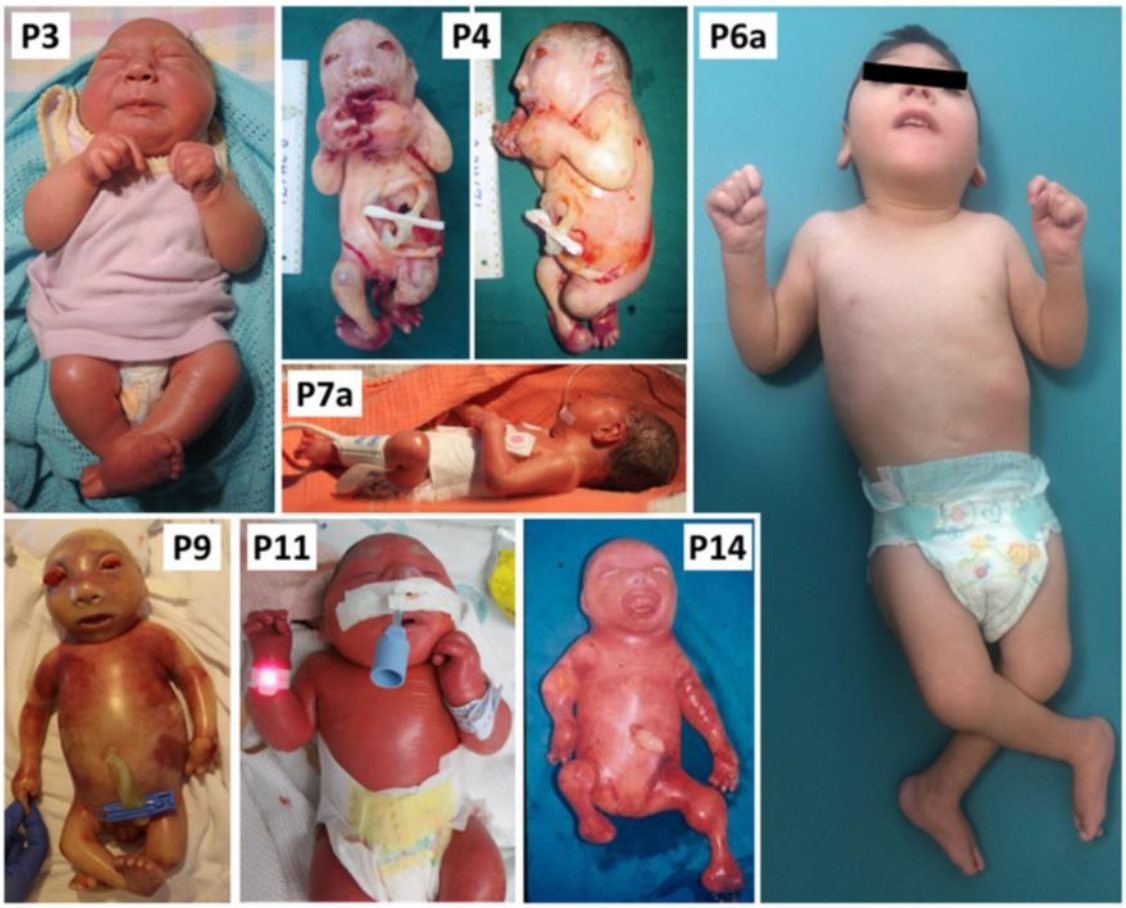


Figure 14: Clinical photographs in selected cases of this cohort.

Clinical photographs documenting the phenotype in selected cases of this cohort. Variable clinical presentation of several patients as newborns (P3, P4, P7, P9; P11; P14) and at the age of 3,5 years (P6a). Patient ID is indicated on the top of the respective photos.

All patients examined showed typical craniofacial abnormalities with microcephaly (-3.4 SD to -6.1 SD for term or near-term newborns), sloping forehead, and micrognathia. A round gaping mouth, low-set and malformed ears, and a short neck were recorded in a large majority of cases. More severe craniofacial abnormalities with frank ocular proptosis, everted lips, and ectropium were less frequent and appeared to correlate with the degree of microcephaly and skin involvement (**Table 1**, **Appendix F Table 1**).

Structural CNS abnormalities were evaluated by ultrasound, MRI or at autopsy. Abnormal gyration was found in 9 of 11 cases and was variable in expression ranging from lissencephaly to

pachygyria and polymicrogyria. Cerebellar hypoplasia, hypoplasia/agenesis of the corpus callosum, and hydrocephalus/enlarged ventricles were identified in almost all cases who had detailed examination. Histopathological CNS findings recorded in one patient (patient 2) included widespread gliosis, hypoplastic corticospinal tracts, reduced number of anterior horn motor neurons, dysplastic thalami, and abnormal brain vessels.

All affected fetuses and infants had some degree of joint contractures, often with a typical pattern of hand and feet posture (**Figure 14**). Swelling or edema of hands and/or feet occurred in 10 of 18 patients (56%). Ichthyosis was recorded in all but two cases was variable in expression. Seven affected individuals showed restrictive dermopathy (39%). Genitourinary abnormalities were reported in 11 of our patients, mostly small kidneys and hypoplastic genitalia. Cleft palate was present in five patients, cataracts were recorded in eight cases. Occasional abnormalities that were recorded included heart defects (patients 9, 12, and 13), sacral agenesis (patient 9), duodenal atresia (patient 7a), underdevelopment of muscles (patient 14), microcornea (patient 12), retinal detachment (patient 10), and preaxial hexadactyly (patient 7a) (**Table 1**, **Appendix F Table 1**).

4.4 Generation of koHEK cells

With the goal of the inactivation of *PHGDH* of wtHEK cells, transfection of wtHEK cells with the constructed *PHGDH* targeting CRISPR nuclease vector was done. Analysis of the percentage of OFP-expressing cells at day 3 post-transfection via flow cytometry revealed a distinct population of up to 72% OFP-expressing cells. To confirm the genome editing of wtHEK cells, a cleavage assay was performed and visualized by a gel imaging system. The analysis of the gel by D1000 ScreenTape assay in combination with the TapeStation system revealed that the percentage of

cleavage was about 65% (**Figure 15C**). These results provided indirect information, that most of the transfected wtHEK cells were edited.

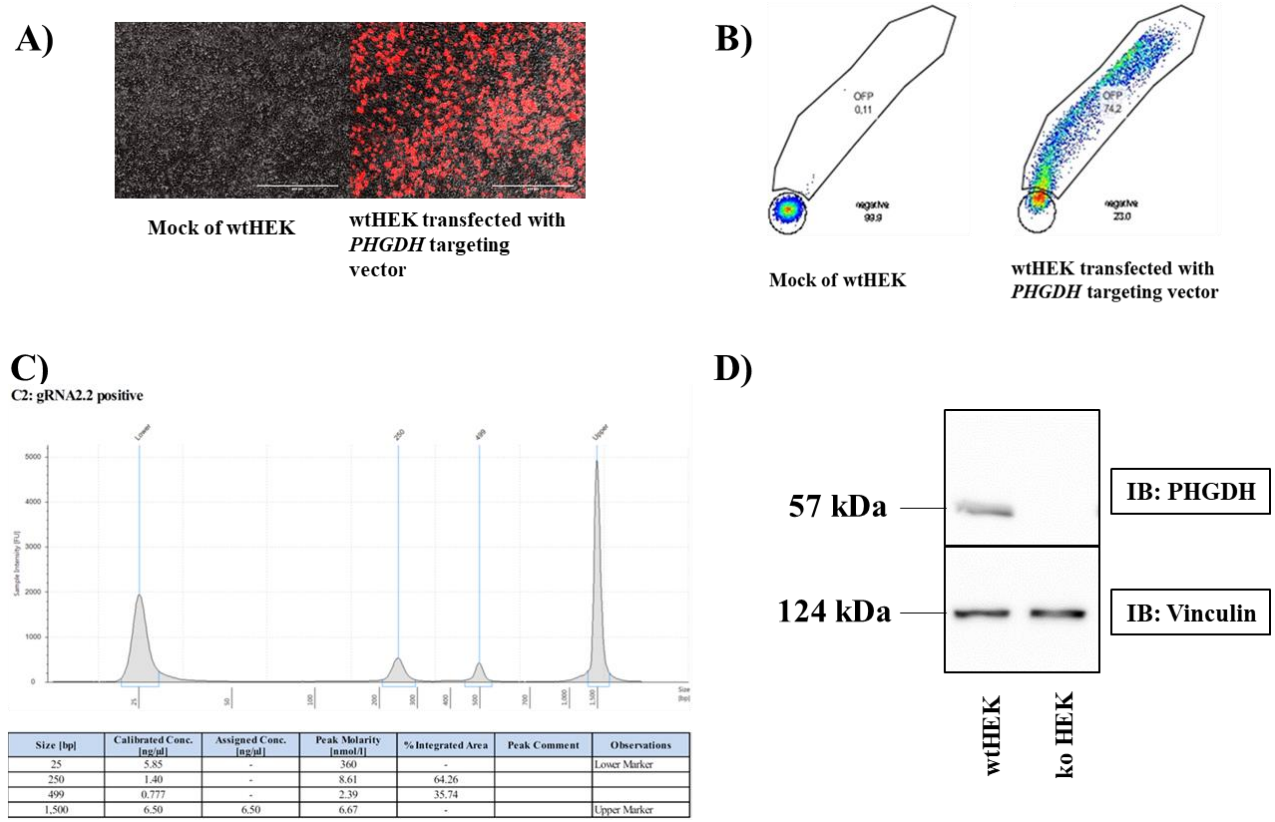


Figure 15: Characterization of koHEK cells.

A) *OFP expression in the transfected wtHEK cells with PHGDH targeting CRISPR nuclease OFP vector, in comparison to mock transfection in wtHEK cells. The scale bar is 400 µm.* ***B)*** *Measurement of transfection efficiency by flow cytometric analysis in each mock and transfected wtHEK cells.* ***C)*** *Gel imaging analysis by TapeStation system to estimate the cleavage efficiency shows the size of the cleaved bands with reference peaks shown at 25 bp and 1500 bp flanking the peaks for cleaved DNA. The cleavage product consists of two relevant bands. The concentration of each DNA fragment was automatically calculated based on the fluorescence intensity. The x-axis in this electropherogram represents the size of the fragments estimated by basepair and the y-axis represents the sample intensity.* ***D)*** *Immunoblotting of PHGDH; while vinculin is equally present in both, wtHEK and koHEK cells, the latter shows the complete absence of the corresponding band of PHGDH. Molecular mass (in kDa) is indicated.*

To confirm the knock-out of *PHGDH*, OFP-positive cell clones were subcultivated and isolated via serial dilution, and the targeted region of *PHGDH* was amplified by PCR. Then the amplicons were cloned into the TOPO-TA vector to obtain single allelic clones for DNA sequencing. About eight single allelic TOPO-TA clones were sequenced per individual cell clone. The sequencing

identified cell lines with two deletion alleles (**Appendix E Figure 1**), where the first allele was (c.770_773delCTGC), and the second allele was (c.759_769delGTGTGCCGGGG). Both deletion alleles resulted in premature stop codons at codon 306 and codon 272, respectively, and are expected to be cleared by nonsense-mediated mRNA decay or to produce a truncated non-functional PHGDH protein. The functional knock-out was confirmed by the absence of the corresponding band for PHGDH protein by immunoblotting (**Figure 15D, Appendix E Figure 5**).

The enzyme activity assay, which was used later, confirmed the near-zero values for PHGDH enzymatic activity in these genetically engineered koHEK cells (as shown in both paragraphs 4.5 and 4.6.1 later). In addition, qPCR confirmed the absence of *PHGDH* mRNA expression in these cells (**Appendix E Figure 6**). Based on these results, it was concluded that the disruption of *PHGDH* in wtHEK cells was successful, leading to a complete functional knock-out of *PHGDH* in HEK293 cells (koHEK), thus making these cells an ideal tool for our further studies.

4.5 Transient expression of *PHGDH* variants in koHEK cells

Eleven different mutants as listed in **Table 7**, in addition to the wt construct and empty vector, were introduced to koHEK cells using the transient transfection system. The localization of the recombinant proteins was monitored by the fluorescence of GFP and the expression of the recombinant proteins was in the cytoplasm for all the transgenic cell lines (**Figure 16**, **Appendix E Figure 7**), which represents the physiological intracellular localization of PHGDH.

For better visualization of the localization of the recombinant PHGDH, the nuclei were counterstained with Hoechst 33342, and the comparison between koHEK cells transfected either with *PHGDH*wt construct or with the empty vector of pTagGFP-C was done by three-dimensional imaging. This imaging showed that the expression of GFP was just in the cytoplasm in the koHEK-

tWT, whereas the expression of GFP was in both, nuclei and cytoplasm, in koHEK cells transfected with empty vector (**Figure 16**, **Figure 17**). All the mutated PHGDH proteins showed similar intracellular distribution as koHEK-tWT (**Appendix E Figure 7**).

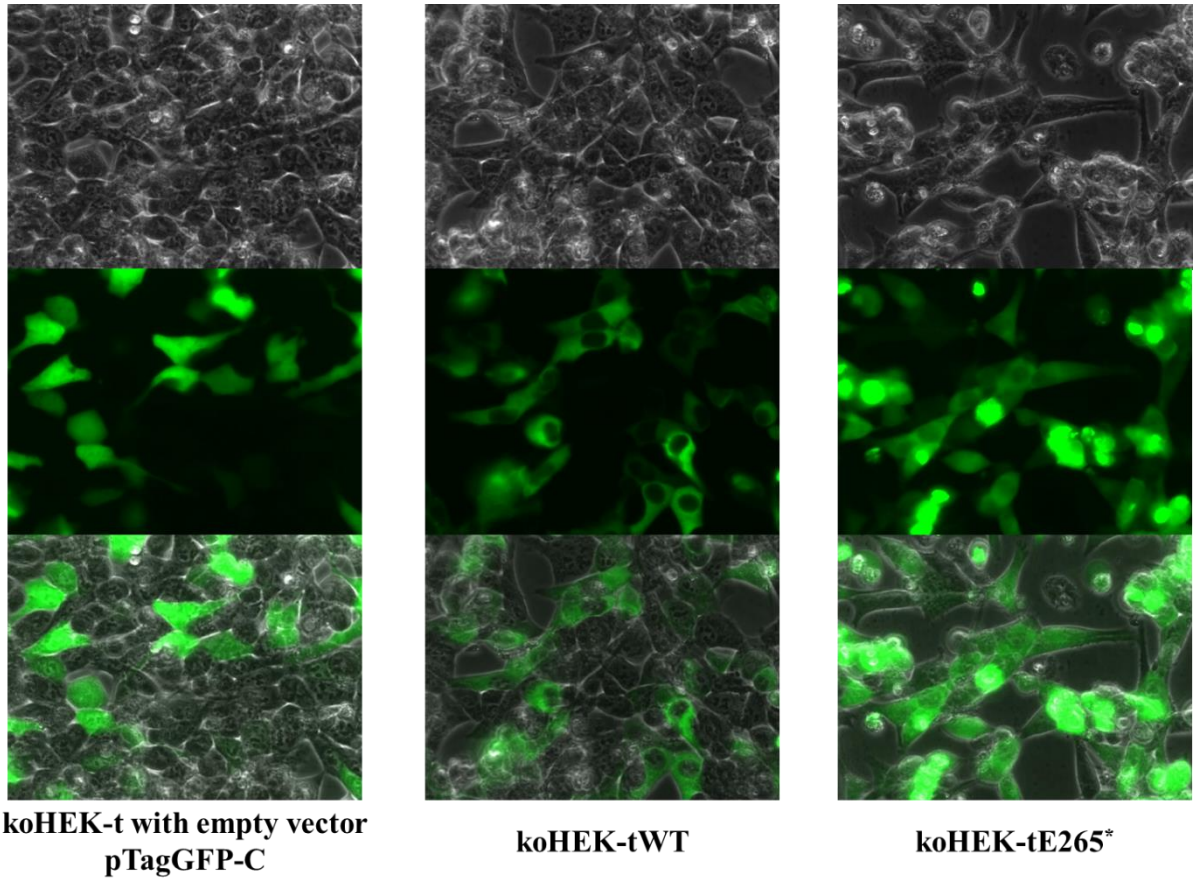


Figure 16: The expression of the recombinant PHGDH in the cytoplasm of transfected koHEK cells.

The top image represents phase contrast, whereas the middle image represents GFP fluorescence using green channel, where the fluorescence microscopy images demonstrate the distribution of GFP fused PHGDH in the cytoplasm of transfected koHEK cells with empty vector, koHEK-tWTand the nonsense mutation c.793G>T in koHEK-tE265. The bottom image represents the merged image.*

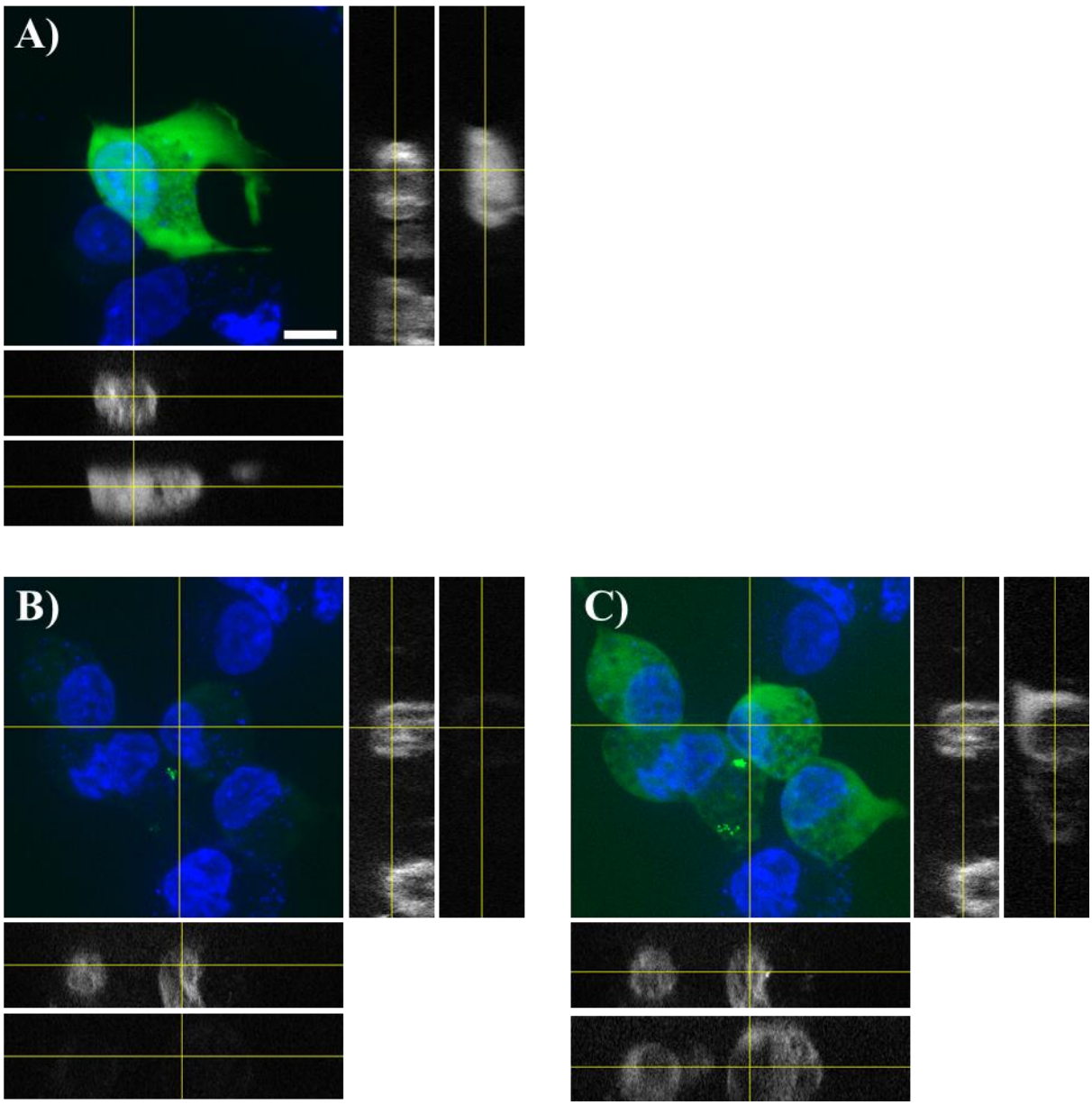


Figure 17: Three- dimensional imaging for better visualization of the localization of the recombinant PHGDH in transfected koHEK.

*High-resolution optical sectioning microscopy, using an Axiovert 200 m equipped with structured illumination by an Apotome2, was used to determine the localization of the PHGDH-GFP. The data is shown here as maximum intensity projections (square image) and orthlogical xz- and yz-sections long the indicated yellow lines. The empty GFP vector in **A)** showed its typical distribution in the cytoplasm as well as in the nucleus. While the same acquisition time settings showed hardly any PHGDH-GFP signal within the koHek cells in **B)**. After a sixfold increase in the exposure time, the PHGDH-GFP signal could be documented within the cytoplasm of the koHek cells. As the orthological sections through the nucleus in **C)** show, the nucleus is not colocalizing with the GFP signal, indicating the cytoplasmic localization of the PHGDH-GFP. The sixfold increase in the exposure time also indicates that the GFP alone is six times more abundant than the PHGDH-GFP fusion within the cells.*

.

Subsequently, PHGDH protein expression as well as enzymatic activity were measured in these transgenic cell lines.

All the transgenic cell lines expressing the mutated PHGDH transiently, showed a significant reduction in the enzyme activity compared to koHEK cells transfected with the wt construct (koHEK-tWT), except two cell lines (koHEK-tE265E, koHEK-tE265K) (**Table 7**, **Figure 18**). The transgenic cell line koHEK-tE265E, which expresses the silent mutation (c.795G>A, p.Glu265Glu) as expected exhibited enzyme activity in the range of the normal control (calculated relative enzyme activity 72%, not statistically significant). Surprisingly, cell line koHEK-tE265K expressing the mutant (c.793G>A, p.Glu265Lys) showed an enzymatic activity similar to that of koHEK-tWT (calculated relative enzyme activity 87%, not statistically significant). This variant was previously recognized as causing NLS mutation (Acuna-Hidalgo *et al.*, 2014) (**Table 7**). Endogenous PHGDH activity in non-transfected wtHEK cells was only slightly and not significantly lower than the activity in the koHEK-tWT cells, thus indicating that the expression of the construct was efficient and in the range of normal PHGDH expression in these cells (**Figure 18**).

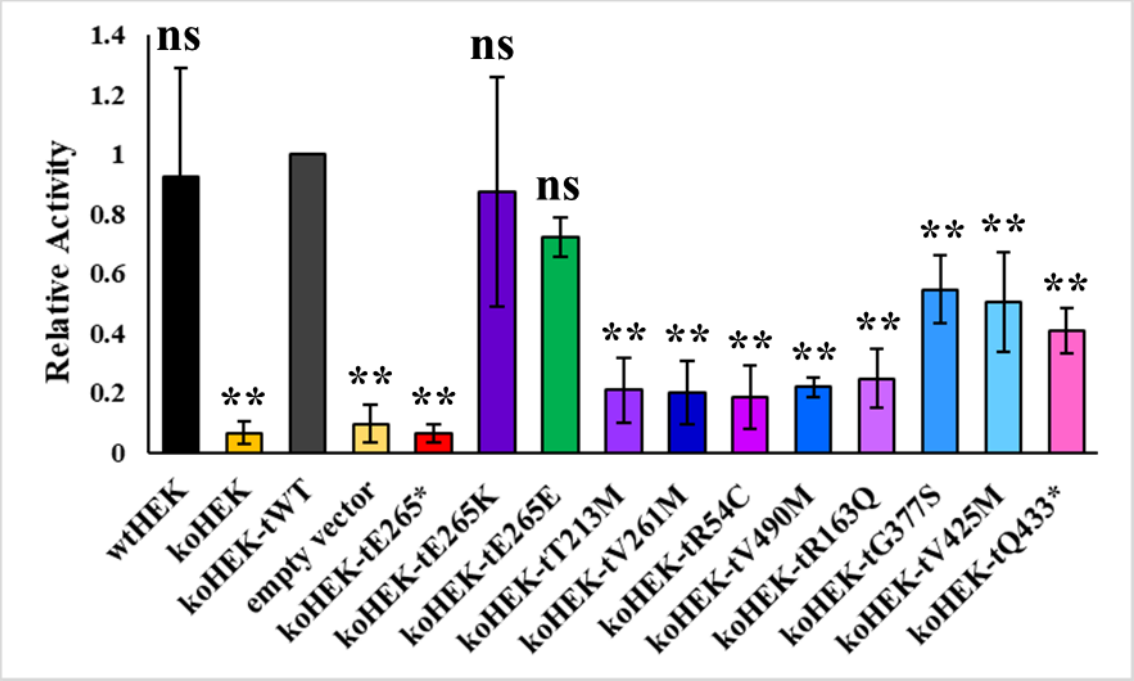


Figure 18: Relative activity of PHGDH activity of variants expressed transiently in koHEK cells.

*Enzymatic activities of mutant constructs relative to the PHGDH activity in koHEK-tWT cells are shown. In addition, empty vector, koHEK, wtHEK cells without transfection were included also and compared to koHEK-tWT. The color blue in its various shades indicates variants observed in non-lethal PHGDH deficiency, whereas the violet color in its various shades indicates variants observed in NLS. Red color indicates a non-natural nonsense mutant as a negative control and green color indicates a non-natural silent mutant as a positive control. Statistical analysis was done by GraphPad InStat using one-way Analysis of Variance (ANOVA) and Dunnett Multiple Comparisons test **p≤0.01, Data represent the mean of four independent experiments n=4. ns indicates not significant when p ≥0.5.*

The cell line expressing a non-natural nonsense mutation (koHEK-tE265*) that served as a negative control showed about 6% relative activity compared to the koHEK-tWT, which was in the same range as the activity measured in the knock-out cells (koHEK), thus confirming the expected loss of function. Three transgenic cell lines expressing mutants observed in NLS showed residual enzyme activities ranging between 18 and 24% (koHEK-tR54C, koHEK-tT213M, koHEK-tR163Q) (**Table 7**, **Figure 18**). Conversely, two transgenic cell lines (koHEK-tG377S and koHEK-tV425M) expressing mutants observed in SBDNL revealed higher levels of relative enzymatic activity ranging between 50% and 54% compared to koHEK-tWT, respectively.

However, two transgenic cell lines expressing mutants observed in SBDNL (koHEK-tV261M and koHEK-tV490M) showed low relative enzyme activities similar to that of NLS-associated cell lines (**Table 7**, **Figure 18**).

Quantitating Western blots of PHGDH protein from wtHEK, koHEK and the eleven transgenic cell lines showed significant reduction in PHGDH expression in eight of the transgenic cell lines when compared to the expression of PHGDH in koHEK-tWT (**Figure 19**, **Appendix E Figure 8 (A, B, C, and D), Table 7**). Reduction in protein expression was moderate in five of them and the relative PHGDH protein expression levels were rather in the range of natural PHGDH expression in wtHEK cells, possibly representing differences in the transfection efficiency.

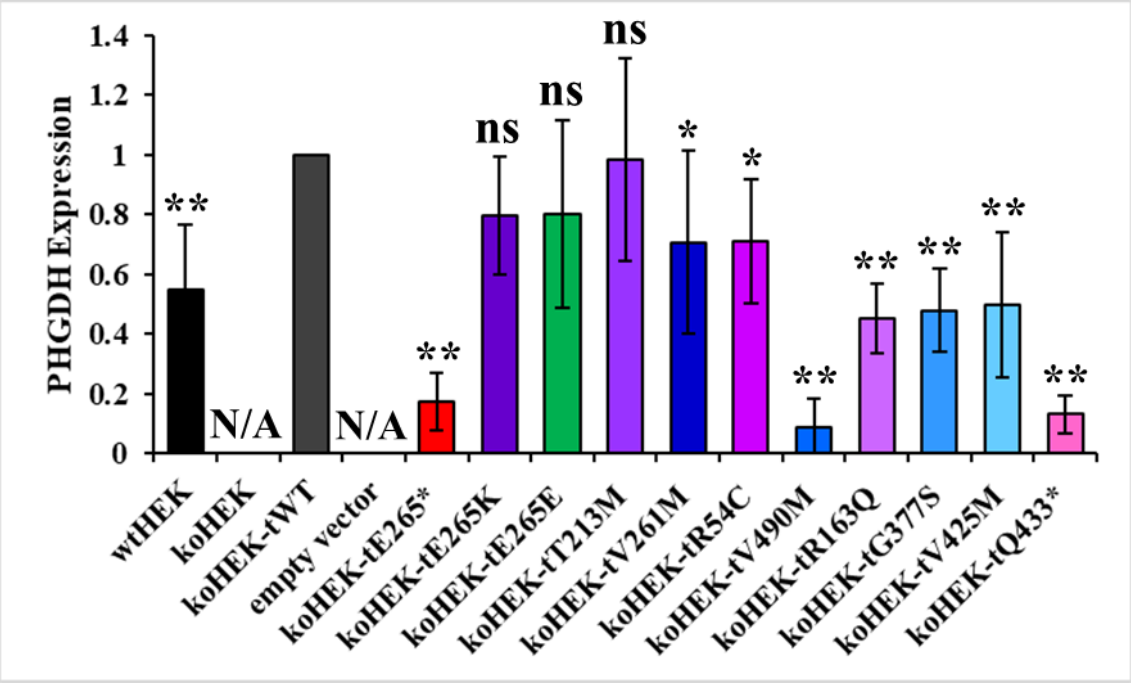


Figure 19: Relative expression of PHGDH in koHEK cells transfected transiently with different constructs.

*Normalization was done to koHEK-tWT, which overexpress wild type of PHGDH. Equal loading was confirmed by immunoblotting for vinculin (original blot images are shown in **Appendix E Figure 8**). The color blue in its various shades indicates variants observed in non-lethal PHGDH deficiency, whereas the violet color in its various shades indicates variants observed in NLS. Red color indicates a non-natural nonsense mutant as a negative*

*control and green color indicates a non-natural silent mutant as a positive control. Statistical analysis was done by GraphPad InStat using one-way Analysis of Variance (ANOVA) and Dunnett Multiple Comparisons test **p≤0.01, * p≤0.05, Data represent the mean of four independent experiments n=4. N/A indicates not available. ns indicates not significant when p≥0.05*.

Unexpectedly, the koHEK-tV490M cell line expressing an SBDNL-associated missense mutant exhibited the lowest protein expression value (8% relative to the overexpressed PHGDH wt in koHEK-tWT) (**Table 7**). This significantly decreased expression might be related to a negative effect of the V490M mutation on the stability of the mutant protein. To validate further this possible effect, proteasome-mediated protein degradation of the expressed protein V490M in koHEK-tV490M was inhibited using MG-132. No clear rescue of mutant protein expression was noticed, thus indicating that the mutant protein is not subjected to increased proteasomal degradation (**Figure 20**, **Appendix E Figure 8 (A, B, C, and D)**). The possibility that a reduced transfection efficiency might stand behind this lower level of mutant protein expression in koHEK-tV490M, was excluded by monitoring the transfection efficiency by fluorescence light microscopy showing nearly similar transfection efficiency as in other cell lines (**Appendix E Figure 7**). It must be considered that this variant V490M might affect the antigenic binding site, but according to the manufacturer, the immunogen sequence used for antibody generation spanning the amino acids from 193-306 (Thermo Fisher Scientific Inc., data sheet on Lot Number: B115626) so that the binding site is located significantly distant from this variant.

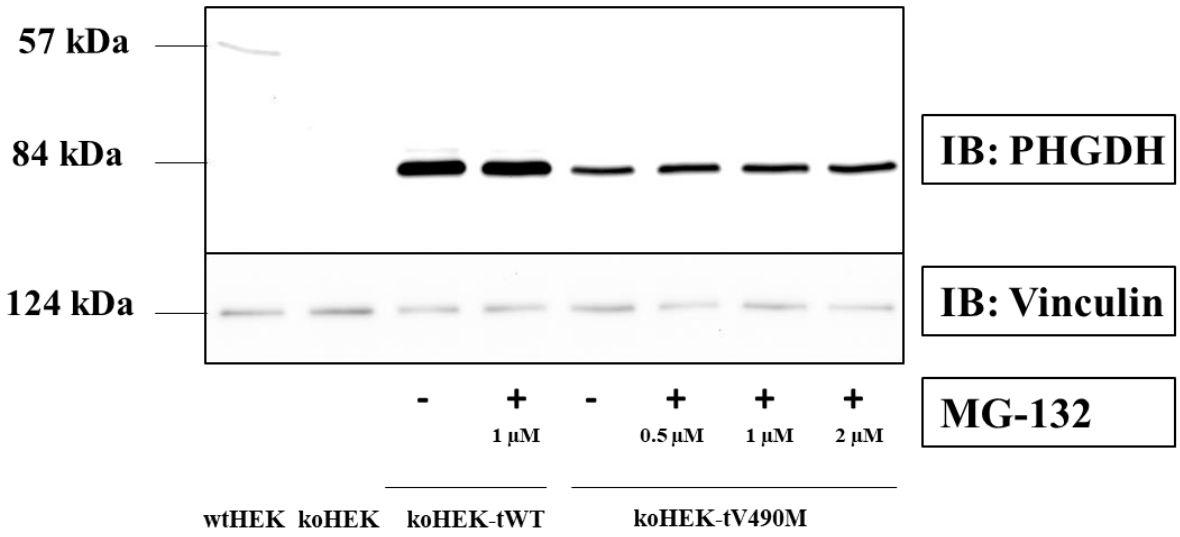


Figure 20: Immunoblotting of* PHGDH *V490M mutant after inhibition of proteasome-mediated protein degradation by MG-132.

To check the possibility of increased proteasomal degradation of the mutated protein expressed in koHEK-tV490M, the proteasome inhibitor MG-132 was applied and protein expression levels were compared to untreated wells and koHEK-tWT. No significant increase in protein expression was noted in cells treated with MG-132. Equal loading was confirmed by immunoblotting of vinculin.

Both truncated forms of PHGDH (p.E265*and p.Q433*) in koHEK-tE265* and koHEK-tQ433*, respectively, showed significant reduction in the expression (relative level of 17% and 13%, respectively, of the overexpressed PHGDH of koHEK-tWT). (**Figure 19**, **Table 7**). This is not surprising, since these truncated protein variants are either not produced due to nonsense-mediated decay of the mRNA or, if produced, may have reduced stability.

No significance difference in the expression of the mutated proteins was observed in three transgenic cell lines (koHEK-tE265K, koHEK-tE265E, and koHEK-tT213M) (**Figure 19**).

Table 7: Results of the studied mutants in two different systems regarding residual enzyme activity and quantification of Western blot results.

Mutant		**Published phenotype associated with variant**	**Reference**	**Relative enzyme activity** (compared to wt PHGDH expressed in koHEK)			**Relative protein expression** (compared to wt PHGDH expressed in koHEK)	
Nucleotide change	**Amino acid change**			**This study (transient overexpression)**	**This study (stable knock-in)**	**Published results with reference**	**transient overexpression**	**stable knock-in**
c.160C>T	p.Arg54Cys	NLS	(Abdelfattah *et al.*, 2020; Acuna-Hidalgo *et al.*, 2014)	18%	N/A	N/A	71%	N/A
c.488G>A	p.Arg163Gln	NLS	(Abdelfattah *et al.*, 2020; Shaheen *et al.*, 2014)	24%	N/A	N/A	45%	N/A
c.638C>T	p.Thr213Met	NLS	(Abdelfattah *et al.*, 2020; Eade *et al.*, 2021)	20%	5%	13% (Eade *et al.*, 2021)[a]	98%	82%
c.781G>A	p.Val261Met	SBDNL	(Coşkun *et al.*, 2009; Tabatabaie *et al.*, 2009)	20%	3%	21% (Tabatabaie *et al.*, 2009)[b]	70%	67%
c.793G>T	p.Glu265*	Non-natural nonsense mutation; negative control	Non natural mutant	6%	2%	N/A	17%	N/A
c.793G>A	p.Glu265Lys	NLS	(Acuna-Hidalgo *et al.*, 2014)	87%	11%	N/A	79%	46%
c.795G>A	p.Glu265Glu	Non-natural silent mutation; positive control	Non natural mutant	72%	98%	N/A	80%	75%

c.1129G>A	p.Gly377Ser	SBDNL	(Benke *et al.*, 2017; Glinton *et al.*, 2018; Tabatabaie *et al.*, 2009; Tabatabaie *et al.*, 2011)	54%	N/A	9-22% (Tabatabaie *et al.*, 2009)[b] 9% (Benke *et al.*, 2017) [b]	47%	N/A
c.1273G>A	p.Val425Met	SBDNL	(Klomp *et al.*, 2000; Kraoua *et al.*, 2013; Meneret *et al.*, 2012)	50%	N/A	3% (Pineda *et al.*, 2000)[b] 58% (Klomp *et al.*, 2000)[a] 9% (Tabatabaie *et al.*, 2009)[b] 45% ((Eade *et al.*, 2021)[a]	49%	N/A
c.1297C>T	p.Gln433*	NLS	(Mattos *et al.*, 2015)	40%	N/A	47% (Eade *et al* 2021) [a]	13%	N/A
c.1468G>A	p.Val490Met	SBDNL	(Eade *et al*., 2021; Klomp *et al.*, 2000; Ni *et al.*, 2019; Pind *et al.*, 2002)	21%	N/A	13-22% (Jaeken *et al.*, 1996)[b] 6-13% (de Koning *et al.*, 1998)[b] 55% (Eade *et al* 2021) [a] 41% (Klomp *et al.*, 2000)[a] 29% (Pind *et al.*, 2002)[a] 28% (Pind *et al.*, 2002)[b] 15% (Tabatabaie *et al.*, 2009)[b]	8%	N/A

Neu-Laxova syndrome (NLS), Serine biosynthesis defect, non-lethal (SBDNL), not available (N/A), in vitro transient expression experiment (a), measurement in patient fibroblasts (b) (individual values for the mutations from Tabatabaie et al. and Pind et al. were calculated from the graph given in the respective figures in the publications.

4.6 Stable expression of *PHGDH* variants in koHEK cells (knock-in *PHGDH* HEK cell lines)

Co-transfection of the koHEK cell line with an AAVS1 TALE nuclease vector and a constructed donor vector (as shown in paragraph 3.2.11.2), either pAAVS1D-PGK-PHGDH-EF1a-copGFP-T2A-Puro or pAAVS1D-CMV-PHGDH-EF1a-copGFP-T2A-Puro was performed to achieve knock-in cell lines with stable expression of wt and mutant PHGDH. This was applied to selected variants (T213M, V261M, E265*, E265K, E265E, wt) and resulted 12 different knock-in *PHGDH* HEK cell lines stably expressing the inserted mutated and wt form of PHGDH. The functionality of the pZT-AAVS1-TALE nuclease pair to cut the *AAVS1* locus and the subsequent success of integration of the provided donor constructs of *PHGDH* was monitored through the constant flouresence of GFP of the cells and its ability to culture and grow in puromycin-containig medium, in addition to the appearance of the expected bands produced by junction PCR in knock-in *PHGDH* HEK cell lines, and absence the band in the negative control in non-transfected koHEK cells (**Appendix E Figure 9**).

Except the cell line transfected with a construct for the nonsense mutation E265* (koHEK-sE265*-CMV), which showed no detectable mutant protein expression, all the transgene *PHGDH* expression from the *AAVS1* locus in the stable transfected *PHGDH* koHEK cells was confirmed by immunoblotting of PHGDH, where the capacity of the generated knock-in *PHGDH* HEK cells to reconstitute PHGDH protein expression was clearly observed with reemergence of a band with the expected size in Western blots, compared to koHEK cells (**Figure 21**).

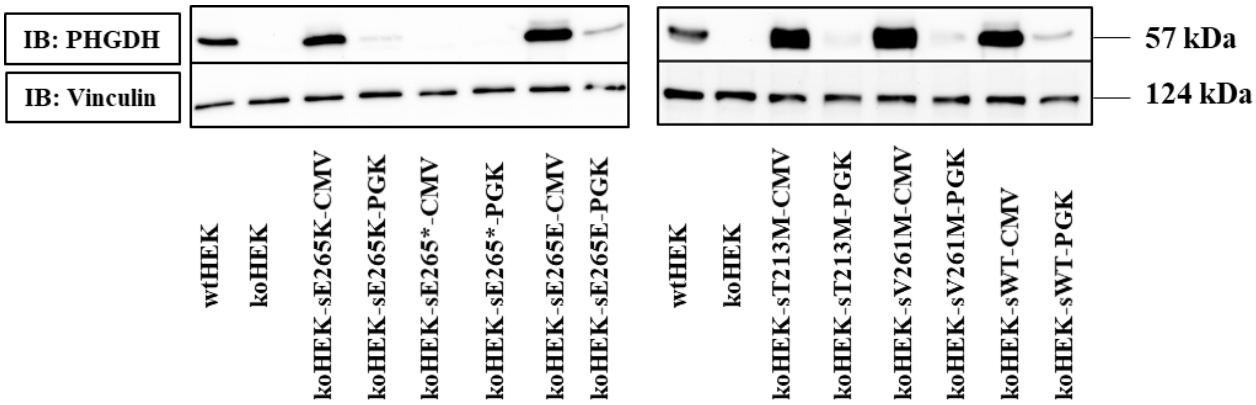


Figure 21: Immunoblotting of PHGDH in wtHEK, koHEK and 12 generated knock-in cell lines.

*It is obvious that koHEK cells and koHEK-sE265*cells fail to express the corresponding band of PHGDH. All the other stable transfected PHGDH knock-in HEK cells with CMV promoter express clear and strong band of PHGDH. In contrast, only weak PHGDH expression was achieved in cell lines transfected with a vector containing the PGK promoter.*

It is obvious that the expression of PHGDH from the *AAVS1* locus generated by the CMV promoter was much stronger than the expression of PHGDH from *AAVS1* locus generated by the PGK promoter. Cell lines with the PGK promoter failed to express the respective PHGDH proteins in a comparable level as the expression of PHGDH from wtHEK cells. The same results were confirmed by qPCR (**Appendix E Figure 6**). Therefore, the experiments described in the following steps were only performed with the cell lines expressing the construct under the CMV promotor, which showed abundant PHGDH protein expression.

4.6.1 PHGDH expression and enzyme activity assay

All stable transgenic cell lines expressing PHGDH variants under the CMV promoter showed a very significant reduction in the enzymatic activity compared to koHEK cells transfected stably with the *PHGDH*wt construct (koHEK-sWT), except one cell line, which expressed the silent mutation (koHEK-sE256E) (**Figure 22**).

The expresed protein E265E from koHEK-sE265E showed a fully normal enzymatic activity (98% relative to the activity measured koHEK-sWT), thus confirming the validity of this construct as a positive control. In contrast, relative enzyme activities of the proteins in the stable transgenic cell lines expressing SBD-associated mutants and the non-natural nonsense mutation E265* (koHEK-sE265*, koHEK-sE265K, koHEK-sV261M and koHEK-sT213M) exhibited significantly reduced relative enzyme activities ranging between 2% and 11% (**Figure 22**, **Table 7**).

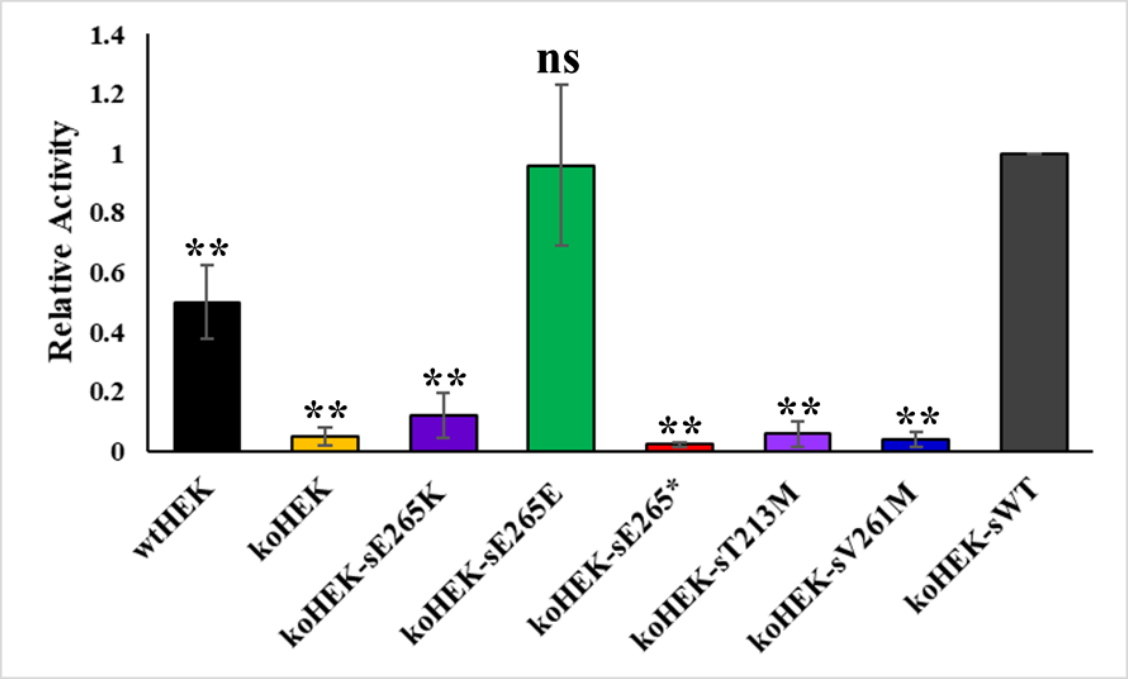


Figure 22: Relative activity of PHGDH variants expressed stably in koHEK (knock-in PHGDH HEK293 cells).

*Enzymatic activities of mutant contructs relative to the PHGDH activity in koHEK-sWT cells are shown. In addition to koHEK, wtHEK cells without transfection were included also and compared to koHEK-sWT. The color blue indicates variant observed in non-lethal PHGDH deficiency, whereas the violet color in its various shades indicates variants observed in NLS. Red color indicates a non-natural nonsense mutant as a negative control and green color indicates a non-natural silent mutant as a positive control. Statistical analysis was done by GraphPad InStat using one-way Analysis of Variance (ANOVA) and Dunnett Multiple Comparisons test **p≤0.01, Data represent the mean of four independent experiments n=4. ns indicates not significant when p≥0.5.*

As expected, the almost undetectable enzymatic activity was measured in the cell line expressing the non-natural nonsense mutation E265*, similar to the koHEK control cell line. Cell lines expressing SBD-associated mutants appeared to have some residual activity, but no significant differences could be determined in this assay between the SBDNL-associated mutant (V261M) and the two NLS-associated mutants (E265K and T213M).

Comparing to the expreesion of koHEK-sWT, most transgenic cell lines (koHEK-sE265E, koHEK-sT213M and koHEK-sV261M) showed PHGDH protein expression levels by quantitating Western blots that were not significantly different from the expression of koHEK-sWT. (**Figure 23**, **Table 7**). One cell line expressing an NLS-associted mutant (koHEK-sE265K) showed significant reduction in the expression of PHGDH (relative expression of 46% compared to the koHEK-sWT cell line).

Since the cell line koHEK-sE265* transfected with a construct for the non-natural nonsense mutation (c.793G>T, p.E265*) showed no band on 7.5% SDS-polyacrylamide gels (**Figure 21**, **Appendix E Figure 10**), the proteins were loaded on 12% SDS-polyacrylamide gels to check the posibility of the presence of the predicted truncated protein (26 kDa) (**Appendix E Figure 5 (A and B)**), but no detectable band was noticed also by this approach, thus indicating that the truncated PHGDH protein E265* is either not produced (clearance of the mutant mRNA by nonsense-mediated decay) or very unstable.

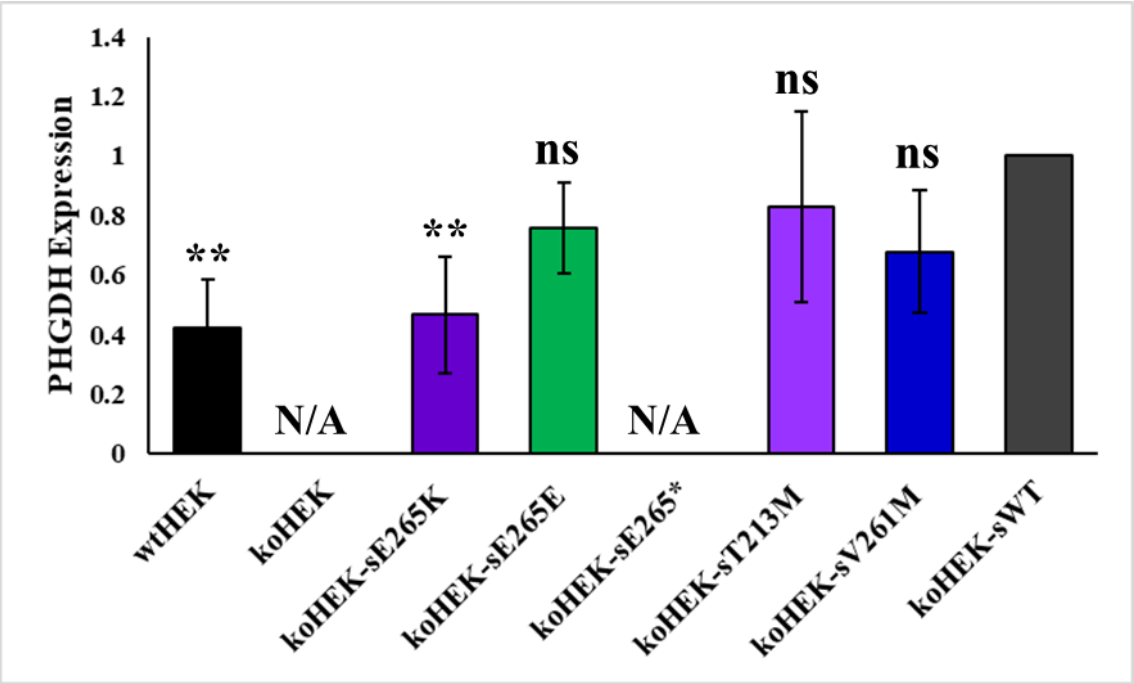


Figure 23: Quantification of PHGDH protein expression in knock-in PHGDH HEK cells.

Protein expression in transgenic cell lines as measured by quantitative Western blotting is expressed as relative value to the level in koHEK-tWTcells. koHEK cells and wtHEK cells were also included. Equal loading was confirmed by immunoblotting for vinculin (shown in ***Appendix E Figure 10****). All these transgenic cell lines expressed the transgene under the CMV promoter. The color blue indicates variant observed in non-lethal PHGDH deficiency, whereas the violet color in its various shades, indicates variants observed in NLS. Green color indicates a non-natural silent mutant as a positive control and no results were available for our negative control nonsense mutation and koHEK cells. Statistical analysis was done by GraphPad InStat using one-way Analysis of Variance (ANOVA) and Dunnett Multiple Comparisons test **p≤0.01, Data represent the mean of four independent experiments n=4. ns indicates not significant when p≥0.5. N/A indicates not available.*

4.6.2 Proliferation and cell viability assays

To assess the biological impact of PHGDH deficiency and disease-associated *PHGDH* mutants on the proliferation and the viability of HEK cells, the generated six different knock-in *PHGDH* HEK-CMV cell lines, as well as wtHEK cells and koHEK cells, were cultured in serine deficient medium. koHEK cells showed a significant impairment of cell proliferation and viability as indicated by BrdU incorporation and the MTS assay (**Figure 24** and **Figure 25**). It could be demonstrated that transgenic expression of wt PHGDH in koHEK-sWT cells was able to reconstitute normal proliferation parameters like wtHEK cells. The same was true for the

transgenic cell line expressing the silent mutation (koHEK-sE265E). Transgenic cell lines expressing disease-associated *PHGDH* mutants showed a significant decrease in their ability to proliferate as measured by BrdU incorporation, except the cell line koHEK-sV261M, showed no significant difference compared to the koHEK-sWT. The V261M mutant is known to be related to non-lethal PHGDH deficiency (**Figure 24**). The lowest values were measured in the cell line expressing the nonsense mutation (koHEK-sE265*) and in the non-transfected koHEK cells.

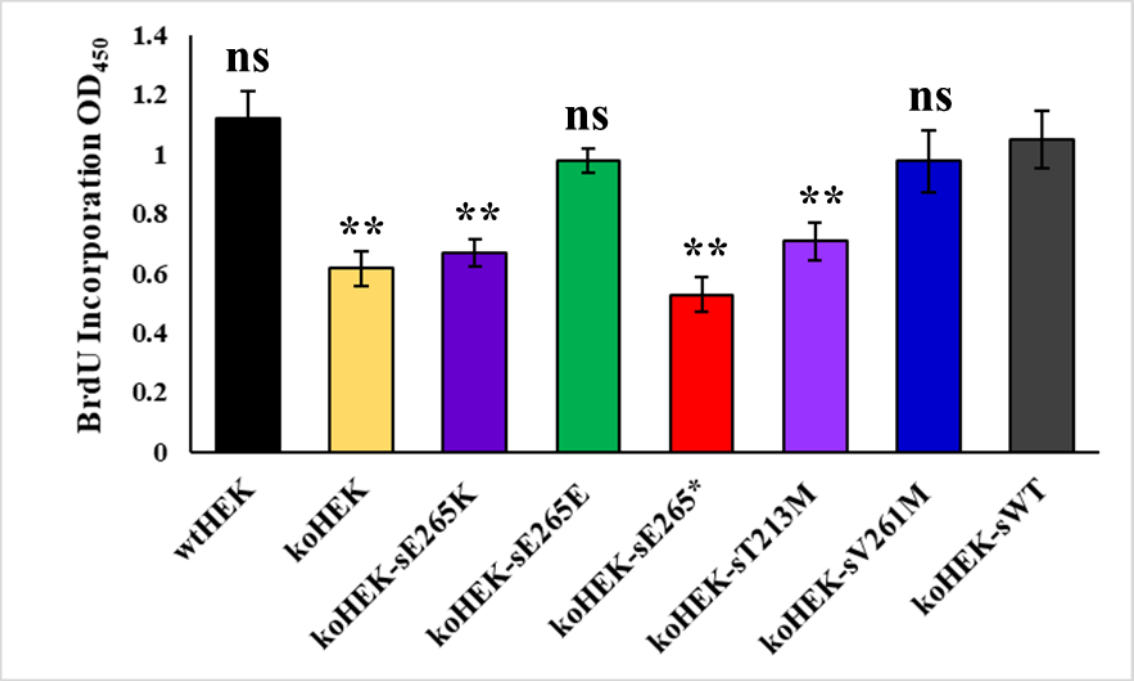


Figure 24: BrdU proliferation assay of knock-in PHGDH HEK293 cells.

*Transgenic cell lines were incubated with BrdU for two hours in serine-deficient medium (MEM supplemented with 1% fetal bovine serum) and BrdU uptake by newly synthesized DNA was measured at 450 nm. koHEK cells and wtHEK cells were included as controls. Values measured in the different cell lines were compared to koHEK-sWT. Statistical analysis was done by GraphPad InStat using one-way ANOVA and Dunnett Multiple Comparisons test **p≤0.01. Data represents mean of three independent experiments n=3. ns indicates not significant and p≥0.05.*

Correspondingly, all *PHGDH*-deficient cell lines showed a significant reduction of metabolic activity in the MTS viability assay. Transgenic expression of wt PHGDH in the koHEK-sWT cell line was able to restore the metabolic activity to the normal range as exhibited by the

wtHEK cell line. The silent mutation expressed in the cell line koHEK-sE265E did not show significant changes in metabolic activity compared to the koHEK-sWT cell line. It is notable that the transgenic cell line koHEK-sV261M, which expresses a SBDNL-associated missense mutation Val261Met showed a higher capacity to restore the metabolic activity of koHEK cell line compared to both NLS-associated mutations in the cell lines koHEK-sE265K and koHEK-sT213M. Again, the lowest values were measured in the cell line expressing the nonsense mutation koHEK-sE265* and in the non-transfected koHEK cells.

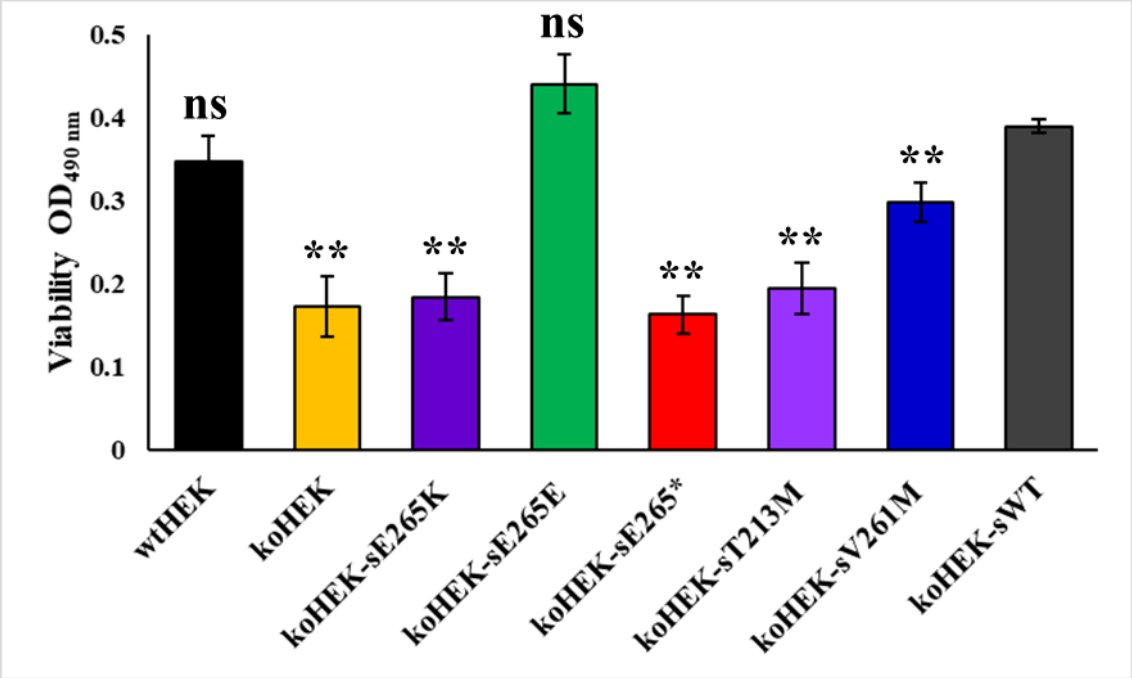


Figure 25: MTS viability assay of knock-in PHGDH HEK293 cells.

*Transgenic cell lines were incubated with MTS for three hours in a serine-deficient medium (MEM supplemented with 1% fetal bovine serum) and cells were lysed, and the absorbance was measured at 490 nm. koHEK cells and wtHEK cells were included as controls. Values measured in the different cell lines were compared to koHEK-sWT. Statistical analysis was done by GraphPad InStat depending on one-way ANOVA and Dunnett Multiple Comparisons test **p≤0.01. Data represent the mean of three independent experiments n=4. ns indicates not significant and p≥0.05.*

5 Discussion

5.1 Mutational spectrum in SBDs and genotype-phenotype correlations

The present cohort (15 unrelated families, 14 of them with NLS and one with a very severe form of PSAT1 deficiency) increased the number of molecularly characterized unrelated families affected by NLS to a total of 35 (Abdelfattah *et al.*, 2020; Acuna-Hidalgo *et al.*, 2014; Bourque *et al.*, 2019; El-Hattab *et al.*, 2016; Kapoor *et al.*, 2021; Mattos *et al.*, 2015; Shaheen *et al.*, 2014; Takeichi *et al.*, 2018). In the combined cohort of previously reported cases and the ones investigated in this study, a *PSAT1* defect turned out to be the most common cause of NLS (accounting for 18 out of 35 unrelated families; 51%), closely followed by *PHGDH* (16/35; 46%).

This distribution may be affected by population selection, since in the present and previously published NLS cases, origin from the Middle East was predominating with at least one obvious *PSAT1* founder allele (c.296C>T, p.Ala99Val). This missense change has been identified in four unrelated families of the current cohort and was previously reported in another five unrelated families (Acuna-Hidalgo *et al.*, 2014; El-Hattab *et al.*, 2016). Notably, *PHGDH* variations are more prevalent in SBDNL (**Figure 13**, **Appendix F Table 2**). For *PSPH* there is still only one NLS family (3%) with a presumed causative variant (Acuna-Hidalgo *et al.*, 2014). Therefore, this gene is still awaiting confirmation of its association with NLS. The underlying genetic defect in one of the two families previously reported as negative for mutations in the three serine biosynthesis pathway genes (family 4, previously reported as family 11 by Acuna-Hidalgo *et al*. (Acuna-Hidalgo *et al.*, 2014)) was solved in this work. There is only one remaining family from that previous study with a tentative clinical diagnosis of NLS and an

unidentified genetic cause. However, the phenotype in that family did not unambiguously support the diagnosis of NLS after a careful review of the clinical data from this family. Three other families with a tentative diagnosis of NLS suspected by the referring clinical partners, the clinical classification remained vague due to features that were inconsistent with NLS or due to a lack of more detailed clinical information. These cases were therefore excluded from the cohort for this study (the studied cohort was restricted to 15 families). Consequently, the hypothesis of possible further genetic heterogeneity of NLS (Acuna-Hidalgo *et al.*, 2014) cannot be further supported. The full genotype could not be fully resolved in all patients: In patient 11, it was possible to identify only the disease-causing variant on one allele, but the presence of the recurrent NLS-associated *PSAT1* variant (c.296C>T) in the heterozygous state is strongly suggesting that this is another case of *PSAT1*-associated NLS, very likely with a different unidentified *PSAT1* mutation on the second allele. The poor quality and quantity of the patient's DNA extracted from the FFPE tissue samples and the unavailability of a paternal DNA sample prohibited a full screening for the paternal variant.

The discovery that genes encoding enzymes of the L-serine biosynthesis pathway are mutated not only in SBDNL but also in NLS, suggested that the latter phenotype represents the severe end of the spectrum of SBDs and that the severity of the phenotype corresponds to the level of functional incapacity of the mutant enzymes (Acuna-Hidalgo *et al.*, 2014; Shaheen *et al.*, 2014). A functional proof for this hypothesis has not been provided. However, it was noted that SBDNL-associated and NLS-associated variants in those genes hardly overlap. The present cohort of 14 novel NLS families further expanded the spectrum of NLS-associated *PSAT1* and *PHGDH* variants (7 and 2 novel variants, respectively), and genotype analysis confirmed the non-overlap with genetic changes causing SBDNL (**Figure 13**, **Appendix F Table 2**, and **Appendix F Table 3**). Family 6 was observed with an intermediate phenotype (between a mild

form of NLS and a very severe expression of SBDNL). This observation supported the view that there is a continuum of phenotypic expression from the most severe prenatally lethal forms (as seen in patients 12, and 14), moderate NLS expression allowing a short period of postnatal survival (patients 3, and 7a) to a very severe non-lethal type of serine deficiency (patients 6a, and 6b) (**Table 1**). It is known that serine plays a vital role in cellular proliferation and the development of CNS and other organs, and *de novo* biosynthesis of L-serine from 3-PG is an essential source of L-serine in mammals (Furuya, 2008). The non-lethal types of serine biosynthesis defects share with NLS the prominent impact on brain development and function, but the degrees of microcephaly and brain dysfunction are extremely variable. The occurrence of other organ malformations in NLS might indicate that more severe incapacity of *de novo* serine biosynthesis affects also early embryonic development.

The severity of the disease did not correlate with the identity of the mutated gene. Less severe expression of NLS was instead observed with variants in either *PHGDH* (e.g., patient 3 with the homozygous mutation c.638C>T) or *PSAT1* (e.g., patient 7a with the compound heterozygous changes c.181C>T and c.296C>T) (**Appendix F Table 1**). Biallelic nonsense, frameshift, or splice site changes, predicting complete loss-of-function of *PSAT1* or *PHGDH* gene products, have never been reported in SBDNL and only one case of NLS (Mattos *et al.*, 2015) (**Appendix F Table 2**). Notably, this affected fetus had a most severe expression of NLS. In the current cohort, a homozygous splice acceptor change (c.870-1G>T) in the *PSAT1* gene was observed in patient 14 with a very severe expression of NLS (**Figure 14**). Notably, the same splice site change was also found in compound heterozygosity with two different missense variants (p.Ala99Val and p.Glu155Gln) in two liveborn patients from unrelated families in the current cohort (patients 10, and 13) (**Appendix F Table 3**) These findings further supported the view that a complete or near-complete loss of enzymatic activity leads to the most severe

expression of NLS. Such cases are presumed to end with early fetal loss and are likely to remain undiagnosed. Consistently, there is one report on a homozygous truncating *PHGDH* variant, c.1030C>T, p.Arg344*, in a case of fetal loss due to non-immune hydrops fetalis (Monies *et al.*, 2019). In the presence of one allele with a variant causing complete loss of function, the severity of expression appears to be dictated by the nature of the change on the second allele. The phenotypic expression in such cases and those with homozygous variants allow a preliminary empirical classification of *PSAT1* and *PHGDH* variants according to their impact on protein (enzymatic) function (**Appendix F Table 2, Appendix F Table 3**). Notably, the *PSAT1* allele c.296C>T, p.Ala99Val, has repeatedly been observed with a somewhat milder expression of NLS, when this change was in the homozygous or a compound heterozygous state (Acuna-Hidalgo *et al.*, 2014). Consistently, the same homozygous variant has recently been reported in a patient that died at 9 weeks of life and was classified as an intermediate phenotype between NLS and infantile serine biosynthesis defect (El-Hattab *et al.*, 2016). Comparison of all cases with the recurrent *PSAT1* missense change p.Ala99Val in a homozygous state also showed some variability, but the phenotypic spectrum appears to be shifted towards a milder expression (all cases were liveborn at term or near term and were lacking frank proptosis or restrictive dermopathy). Accordingly, the *PSAT1* variant c.129T>G, p.Ser43Arg, which was identified in family 6 with the intermediate clinical expression has been published before with a similar phenotype of SBD with severe prenatal onset microcephaly (Brassier *et al.*, 2016). As it seems that gradual differences in residual enzymatic activity may have a great influence on the phenotype, we anticipate quite tight genotype-phenotype correlations, but many more observations are necessary to corroborate variant-specific phenotype associations. Additional unidentified genetic modifiers or non-genetic factors may also play a role in determining the severity of expression. These considerations underscore the need for functional studies to assess

the residual enzymatic activity of mutants associated with variable expressions of SBD. A very recently published study using a quantitative, yeast-based growth assay could demonstrate differences in the functional impact of *PSAT1* variants, which were in agreement with phenotype annotations (Sirr *et al.*, 2020). The authors reported normalized growth estimates that were lowest for some NLS-associated variants, such as p.Ser179Leu and p.Cys245Arg (<0.10), intermediate for the variant p.Ala99Val (0.60) associated with a less severe expression of NLS, and mildly impaired for p.Ser43Arg (0.75) reported with a severe form of non-lethal PSAT1 deficiency. The highest level, but still significantly below the normal, was calculated for p.Asp100Ala, a variant known to be associated with non-lethal PSAT1 deficiency (Sirr *et al.*, 2020).

When comparing the localization of mutations observed in SBDNL and NLS, it is noticed that the variants were distributed along the *PHGDH* gene (**Figure 13A**), but interestingly more than half of the variants identified with a non-lethal PHGDH deficiency disorder resided in the regulatory domain of PHGDH. In contrast, half of the NLS-associated variants were in the nucleotide-binding domain of PHGDH (**Figure 13A)**. It was previously suggested that alteration of the C-terminal domain of mammalian PHGDH had less impact on the enzymatic activity (Klomp *et al.*, 2000). It was also noted that the presumed regulatory domain had poor conservation across species (Achouri *et al.*, 1997). In the latter study, it was shown that the removal of the carboxyl-terminal 209 amino acids from the *R. norvegicus* enzyme lowered but did not abolish the enzyme activity (Achouri *et al.*, 1997). However, differential effects of variants in those genes are not strictly related to the affected protein domains, they may also occur within the same domain, even with mutations at closely neighboring codons. This is exemplified by *PSAT1* mutations p.Asp100Ala and p.Ala99Val, which are known to cause non-lethal PSAT1 deficiency and NLS, respectively (**Figure 13B, Appendix E Figure 4**).

In conclusion, this cohort together with a careful review of previously reported cases showed that pathogenic *PHGDH* and *PSAT1* variants lead to a spectrum of human disorders and that NLS *per se* has varying degrees of phenotypic expression likely representing the extreme end of a continuum. The assumption that the individual residual enzyme activity of mutant proteins might be the major determinant of the phenotypic variability (Abdelfattah *et al.*, 2020; Acuna-Hidalgo *et al.*, 2014; Shaheen *et al.*, 2014), was the motivation for performing functional experiments to explore in detail the functional basis of the phenotypic variability.

5.2 Functional analysis of *PHGDH* mutants

The serine biosynthesis pathway is a key element contributing to a wide spectrum of SBD (Abdelfattah *et al.*, 2020; El-Hattab *et al.*, 2016). PHGDH is the first rate-limiting enzyme in the serine biosynthesis pathway. *PHGDH* mutations occur with the almost equal frequency as *PSAT1* mutations in NLS, but *PHGDH* is more prevalent in SBDNL. To compare NLS-associated and SBDNL-associated mutants, the functional experiments were focused on PHGDH in this thesis work. For a few *PHGDH* mutants associated with non-lethal types of serine biosynthesis defects, enzyme activity measurements have been reported with a considerable residual enzymatic activity in the range of 3-58% of the normal (Klomp *et al.*, 2000; Pind *et al.*, 2002; Tabatabaie *et al.*, 2011) (**Table 7**). However, the materials and assays used in those studies are hardly comparable. Previous functional studies on SBDNL-associated *PHGDH* mutants used various approaches and most of them included only a single or very few variants, in some of them, positive and negative controls were missing or inappropriate (Pineda *et al.*, 2000). Some studies were done on patient fibroblasts cells (Benke *et al.*, 2017; de Koning *et al.*, 1998; Jaeken *et al.*, 1996; Tabatabaie *et al.*, 2011), while others used in vitro

overexpression systems (Klomp *et al.*, 2000; Pind *et al.*, 2002). Notably, the results of residual enzyme activities for mutants that were evaluated in multiple studies (e.g., V425M, V490M) showed a huge variability between the different studies (**Table 7**). As shown in Table 7, there is an obvious discrepancy in the calculated residual enzyme activity levels between the measurements in patient's fibroblasts (de Koning *et al.*, 1998; Jaeken *et al.*, 1996) compared to studies using molecular cloning and transient overexpression of mutant enzyme proteins in wild type-expressing cells (Klomp *et al.*, 2000; Pind *et al.*, 2002) with substantially higher enzyme activity levels determined with the latter method (V425M: 3-9% in patient fibroblasts vs. 45-58% in overexpression experiments; V490M: 6-22% in patient fibroblasts vs. 41-55% in overexpression experiments, **Table 7**). One major reason for the higher levels measured in overexpression experiments is likely the intrinsic enzyme expression of the cells. In all previous *in vitro* overexpression studies cells having intrinsic enzyme expression were used. Only in one study, measurements of the enzyme activity of the mutant V490M in patient fibroblasts and in overexpression experiments yielded similar results (28% vs. 29%) (Pind *et al.*, 2002). Some variability was also discovered when patient cells from members of the same family were investigated: Jaeken *et al*. observed 13% vs. 22% of normal PHGDH activity in fibroblast from brothers with the V490M mutation (Jaeken *et al.*, 1996), whereas Klomp *et al.* measured 6% vs. 13% of the normal PHGDH activity in overexpressed V490M in one of both brothers of Jaeken *et al.* patients (Klomp *et al.*, 2000).

As shown above, the transient overexpression studies can harbor many pitfalls. The variation in the transfection efficiency is one of them. In the current study the impact of varying transfection efficiencies is noticeable by higher standard deviations within the activity measurement when compared to the same mutation in the stable transfection approach (**Figure**

18, **Figure 22**). Pind *et al.* tried to solve that problem by co-transfection of a β-galactosidase coding vector and normalizing to the β-galactosidase activity (Pind *et al.*, 2002). Another problem is the control of the copy number of the vectors per cell. While a transient transfection systems depend more or less on a high copy number of the overexpressed variant, specially when wild type cells were used as in several previous studies (Eade *et al.*, 2021; Klomp *et al.*, 2000; Pind *et al.*, 2002), to measure the effects of mutations on the PHGDH activity. This could lead to an overestimation of the mutant activity through the fact that the intrinsic enzyme expression has a higher relative contribution to the overall activity in cells transfected with the mutant compared to those transfected with the wild type and through the mechanism that overexpression of the wild type is likely to reach a saturation level for the metabolic throughput at a lower abundance of the protein than in case of the mutant enzyme. In our study we overcame the confounding effects of intrinsic enzyme activity using knock-out cells for the expression of the mutants of interest. Nevertheless, we still calculated higher enzyme activities for the mutants in the transient expression system than in the more reliable stable transfection approach (T213M: 20% vs. 5%; V261M: 20% vs. 3%; E265*: 6% vs. 2%). Several studies have shown negative effects of harsh overexpression in various expression systems, ranging form *E. coli*, over *S. pombe* and *C. elegans* to mammalian cell lines as summarized for example by Prelich (Prelich, 2012). Earlier studies did not have the access to knock-out cells or the technology to produce them as in this study. Their approach could only be to overexpress so much mutated PHGDH that the protein served as a dominant negative variant against the still expressed wildtype variant. This can only be achieved when the amount of wild type PHGDH within the cell lysate is overpowered of the mutated form in any activity assay used, which could lead to the adverse effects mentioned above. Another problem for the reproducibility of *in vitro* transient expression experiments with subsequent activity measurements are missing

methological informations. While Pind *et al.* nicely described their cloning strategy and transfection approach down to the amount of used vector and the volume of used Lipofectamine transfection reagent, the crucial information, how many cells in what kind of vessel were transfected, is missing. (Pind *et al.*, 2002).

Over the years, also the methodology has changed drastically. While Jaecken *et al*., Koning *et al*., Pineda *et al.* and Klomp *et al*. were all using the reductions of the absorbance of NADH at 340 nm in their activity assays utilizing the reverse nonphysiological direction of the PDGDH activity (de Koning *et al.*, 1998; Jaeken *et al.*, 1996; Klomp *et al.*, 2000; Pineda *et al.*, 2000). Eade *et al*. and this study are the first to utilize an assay that can detect the activity in the physiological direction (Eade *et al.*, 2021).

Considering those previous results, we decided to use a multimodal approach with different types of transgene expression and different readout systems (enzyme activity measurements and cell proliferation assays) to validate results by different methods and to obtain a more complete picture. The selection of the methods employed in this study also had to reflect the unavailability of patient cell lines from NLS patients. Therefore, the study had to rely on *in vitro* models representing the mutants of interest. To explore the functional incapacity of *PHGDH* mutants, its correlation with either NLS or SBDNL, several variants that were considered representative for either of these disease types were selected from (**Appendix F Table 2**). In addition, two non-natural variants, a silent one and a nonsense variant, were generated and included as positive and negative controls (**Table 7**). Taking advantage of the CRISPR/Cas9 system it was first intended to target *PHGDH* and introduce the selected mutations at the same time by HDR. Unfortunately, all the trials led to various length deletions in most of the generated and sorted cell lines, thus indicating efficient targeting of the *PHGDH*

locus by Cas9, but inefficiency of the HDR with this approach. However, in one of these modified cell lines, two deletion alleles were recognized that led to shifting the ORF of *PHGDH*. The efficient and complete knock-out of *PHGDH* in this cell line (koHEK) was confirmed on the mRNA, protein and enzyme activity levels as shown in the respective sections in the Methods and Results (3.2.12.1 and 4.4). This generated koHEK cell line was an optimal basis for the transfection with selected mutants to measure their functional impact, because the koHEK cells were free from any confounding endogeneous PHGDH background activity.

Taking together the results from the different approaches used in this study, it could be reliably demonstrated that all the tested disease-associated mutants do confer reduced or nearly abolished function. The validity of our *in vitro* expression systems (transient and stable) was documented by the complete absence of PHGDH protein expression and enzyme activity in the genetically engineered koHEK cells that served as basis for the expression of transgenes, the full rescue of protein expression and enzymatic activity in the cells transfected with the WT construct, and the appearance of PHGDH protein in the cell lines transfected with mutant constructs. The validity of this approach was further corroborated by the consistently normal results for a non-natural silent variant E265E (positive control) and the consistent complete loss of function of the non-natural nonsense variant E265* (negative control) in all lines of those experiments. The results for the latter (lack of protein expression and enzymatic activity) is most likely explained by nonsense-mediated mRNA decay, as the transfection efficiency was confirmed by monitoring the fluorescence of GFP in both expression systems (**Figure 16**, **Appendix E Figure 7**). Taken together, we can therefore conclude that our test system was useful and valid to confirm the pathogenic relevance of all the investigated disease-associated variants. This is of particular importance since one of the mutations tested (c.638C>T, p.Thr213Met) was a novel one identified in the first part of the study, for which the formal

classification according to ACMG criteria (Richards *et al.*, 2015) remained "VUS" (PM2, PP2, PP3). With the demonstrated functional incapacity of the mutant its classification changes towards "likely pathogenic" (PS3, PM2, PP2, PP3). Our test system or parts of it could therefore serve for the validation of pathogenicity of novel uncertain variants in the future. Since most SBD-associated variants are missense mutations for which the functional consequences cannot be reliably predicted, the availability of a reliable assay to obtain experimental evidence for pathogenicity is critical for variant classification.

Despite using cells that were deficient of endogenous PHGDH, the calculated residual enzyme activities differed for individual variants between the transient and stable transfection systems. Most disease-associated variants exhibited lower levels of residual enzyme activity in the stable transfection system compared to transient overexpression, e.g., T213M: 5% vs. 20%; V261M: 3% vs. 20%; E265K: 11% vs. 87%. As mentioned above, it has been reported also in previous studies that enzyme activities in transient overexpression systems are higher than in patient fibroblasts (de Koning *et al.*, 1998; Jaeken *et al.*, 1996). This discrepancy between the results of transient and stable transfection systems or patient cells has not been clarified experimentally but might be explained by plasmid copy numbers in the host cells in the transient transfection system. While the overexpressed wildtype enzyme may reach a saturation level, increased copy number might amplify the enzymatic activity of hypo-active mutants and thereby in part compensate for the functional incapacity in the serine biosynthesis pathway. However in the experiments presented here, there was no substantial difference in the mutant protein expression relative to wtHEK cells between the transient overexpression and the stable expression under the CMV promotor (**Appendix E Figure 8**, **Appendix E Figure 10**).

Two variants stood out of the 11 mutants tested in the transient overexpression assay: The novel NLS-associated variant E265K exhibited an enzymatic activity in the range of koHEK-tWT and wtHEK cells, thereby suggesting normal function in this assay (**Figure 18**). However, subsequent analysis in the stable transfection system clearly confirmed a significantly decreased enzymatic activity (11% of the koHEK-sWT control) as well as impaired proliferation and metabolic activity of the cell line expressing this particular mutant (**Figure 22**, **Figure 24**, **Figure 25**, and **Table 7**). These results confirm that the E265K variant strongly reduces the activity of PHGDH and support its pathogenetic role for NLS. The reason for the apparent false negative result in the transient transfection system could not be resolved. This could point at a higher robustness and validity of the stable transfection assay compared to transient overexpression. Specific abnormal properties of this variant that may be circumvented in the transient overexpression system, such as impaired protein stability or defective intracellular trafficking, cannot be excluded.

The variant V490M produced a strongly decreased protein expression in the transient transfection system compared to all other constructs (**Figure 19**, **Appendix E Figure 8**), in spite of the nearly similar transfection effeciency of all the variants (**Appendix E Figure 7**). The cause of reduced expression of this particular mutant could not be resolved. Experiments performed with a proteasome inhibitor did not provide evidence for increased proteasomal degradation of this mutant protein (**Figure 20**). It is also unlikely that the variant affects the antigenic motiv for antibody binding which could explain a lower immunoreactivity in the Western blots, since the manufacturer indicates that the antigenic sequence encompassed amino acids 193 to 306. No experiments were performed in this study to exclude that reduced mRNA stability could be the reason. However, a previous study did not find that this variant altered

mRNA synthesis rate or stability (Klomp *et al.*, 2000). Notably, Pind *et al*. (Pind *et al.*, 2002) reported that this particular variant yielded about 29% of the activity observed for WT enzyme *in vitro* and suggested that the reduction in the activity of the corresponding protein resulted from an increased degradation of the mutant protein. Taken together, the present and previously published findings suggest that these variant exhibits protein instability that is not associated with increased proteasomal degradation. Ni *et al.* suggested that this substitution V490M changes the structure of PHGDH, where it could alter the distance between the betasheet and the neighbouring alpha helix and this beta sheet is located at the binding interface of two PHGDH proteins, the mutation could therefore affect dimerization (Ni *et al.*, 2019). Dimer formation of PHGDH has been postulated to facilitate substrate binding and to maintain the conformation required for PHGDH catalysis and activity (Xu, Qing, Wang, Li, & Lai, 2021).

With the knowledge that NLS type1 and PHGDH deficiency are allelic disorders, the hypothesis was raised that disease severity and the phenotype variability can be explained primarily by the degree in functional incapacity or the residual enzymatic activity of the respective mutant enzyme proteins (Abdelfattah *et al.*, 2020; Acuna-Hidalgo *et al.*, 2014; Shaheen *et al.*, 2014). Indeed, the previously reported enzyme activities for non-lethal PHGDH deficiency were significantly above zero – though quite variable in different test systems – suggesting that these mutants confer moderate impairment of function with a significant level of residual enzymatic activity (Benke *et al.*, 2017; de Koning *et al.*, 1998; Eade *et al.*, 2021; Jaeken *et al.*, 1996; Klomp *et al.*, 2000; Pind *et al.*, 2002; Pineda *et al.*, 2000; Tabatabaie *et al.*, 2009). No results from functional testing of NLS-associated *PHGDH* variants have been published, so far. For *PSAT1* variants, however, Sirr *et al*. conducted a systematic study of many variants in a yeast growth assay. These variants included also variants reported to be associated with NLS and

SBDNL. The reported findings were generally in line with the hypothesis that NLS-associated mutants have stronger impairment of function compared to SBDNL-associated variants, but the assay had a too large range of variation to predict precisely the specific significance for individual variants (Sirr *et al.*, 2020). It is notable, that also for NLS biallelic *bona fide* loss-of-function mutations (e.g., biallelic truncating variants) have not been reported, except for one case with a very severe prenatally lethal phenotype (Mattos *et al.*, 2015). It has been proposed that the complete loss of function of the *de novo* serine biosynthesis pathway is early prenatally lethal, while survival of a child with NLS until term likely requires at least a minimal level of residual function (Acuna-Hidalgo *et al.*, 2014). The aim of this study to prove the existence of residual activity of NLS-associated mutations and to resolve more clearly the functional dichotomy between NLS-associated and SBDNL-associated mutations could be achieved partially.

By all lines of investigation (transient and stable expression in koHEK cells, as well as enzymatic and proliferation / metabolic activity testing) the findings consistently suggested that NLS-associated variants have at least a low level of residual functional capacity: The five NLS-associated variants tested in the transient expression system (R54C, R163Q, T213M, E265K, Q433*), as well as the two NLS-associated variants tested in the stable transfection system, all exhibited levels of enzymatic activity or proliferative and metabolic activity above the level of the loss-of-function E265* variant and the non-transfected koHEK cells in the respective assays (**Figure 18**, **Figure 22**, **Figure 24**, **Figure 25**, and **Table 7**). In the transient transfection system, the NLS-associated variants exhibited a relative activity of 18% or more, whereas the nonsense mutation E265* as a presumed functional null variant was negligible (6%) (**Figure 18**, **Table 7**). These findings therefore support the assumption that most NLS-associated variants are

strongly incapacitating hypomorphs but not functional null alleles. However, in the experiments performed here the difference between NLS-associated variants and koHEK cells did not consistently reach statistical significance. In conclusion, the employed experimental methods unfortunately did not reach the resolution to prove the assumption of minor residual function of NLS-associated mutants.

Regarding the functional difference between NLS-associated and SBDNL-associated mutations, the findings of this multimodal experimental approach were generally in line with the hypothesis of less severe functional impairment of SBDNL-associated variants but overlaps and outliers existed that prevented a picture of strict dichotomy. The small number of variants included particularly in the transfection experiments with stable expression (only two NLS-associated variants and one SBDNL-associated variant) was another limiting factor. Focussing on the variants tested in the transient expression system, it is observed that most variants related to NLS showed enzyme activities around 20%, whereas two of the three SBDNL-associated variants (G377S, V425M, V490M) located in the putative regulatory domain of PHGDH exhibited relative enzymatic activity around 50% (**Table 7**). The value measured for V425M are consistent with the results previously reported by Klomp *et al*. (Klomp *et al.*, 2000), while other studies reported an activity of less than 10% of normal for this variant (Pineda *et al.*, 2000; Tabatabaie *et al.*, 2009) (**Table 7**). This illustrates the lack of comparability of results gained from different methodological approaches. The third mutant affecting the regulatory domain, V490M, showed a greater decrease in enzyme activity more in the range of the aforementioned NLS-associated mutants, with relative activity of 21% compared to the activity of transgenic koHEK-tWT cell line. This variant has recurrently been observed in patients with non-lethal PHGDH deficiency (Klomp *et al.*, 2000; Pind *et al.*, 2002), but it was also reported once in a

compound-heterozygous state in a patient classified as NLS (Ni *et al.*, 2019). This might suggest that this variant may have intermediate functional and phenotypic impact between NLS and SBDNL. The results of the enzymatic activity measured for V490M is in the same range as the results of Pind *et al*. (Pind *et al.*, 2002), where they reported that this variant yielded about 29% of the activity observed for WT enzyme *in vitro*. In a study relating heterozygous PHGDH variants to late-onset macular telangiectasia type 2 a relative enzyme activity of about 55% as was reported (Eade *et al.*, 2021) (**Table 7**). The particular phenomenon of low protein expression for this variant has been discussed above. The experiments performed in the stable transfection system can only be considered preliminary regarding the difference between NLS-associated and SBDNL-associated mutants, because of the limited number of mutations included. Nevertheless, it is notable that in particular the proliferation and metabolic activity assays (BrdU and MTS assays) in the cultured transgenic cell lines very nicely displayed the difference between the strong effect of the two NLS-associated mutations compared to a mild effect of the SBDNL-associated variant studied (**Figure 24**, **Figure 25**). This assay on living cells cultured in a serine-depleted medium probably reflects the strong biological impact of severely incapacitating mutations on basic mechanisms of cell growth and proliferation. Such impact is particularly evident in NLS, where affected individuals typically show severe growth retardation from early intrauterine life and hypotrophy of many organs, especially but not exclusively the brain (Cavole *et al.*, 2020). In contrast, enzymatic activity measurements in cells with stable expression of disease-associated mutants were not consistent with the differences recorded in the proliferation and viability assays, as the levels of relative enzymatic activity of the two NLS-associated variants (T213M and E265K) were not distinct from the one of the SBDNL-associated variant V261M.

The proliferation and metabolic activity assays possibly do not have the discriminatory power of enzyme activity measurements, as they depend on multiple variables, but this approach may be closer to the *in vivo* situation. Notably, even koHEK cells were able to proliferate in the absence of serine, but showed very significant reduction in the prolifiration compared to the transgenic koHEK-sWT cell line (**Figure 24**, **Figure 25**). Serine supply in these cells under external serine depletion might stem from alternative routes, e.g., conversion from glycine by the reversible action of serine hydroxymethyl transferase (Ou, Wang, Jiang, Zheng, & Gu, 2015; Wang *et al.*, 2013). Interestingly, the transgenic silent mutation koHEK-sE265E and the transgenic wild type koHEK-sWT apparently improved the proliferation and viability of koHEK cell line compared to wtHEK cells. This may be a similar effect as the one by which PHGDH amplification is driving tumor cell growths, as discussed above (Locasale *et al.*, 2011; Pollari *et al.*, 2011). Several previous studies have addressed the consequences of PHGDH deficiency in proliferating cells. For example, Vandekeere *et al*. showed that PHGDH inhibition impairs heme synthesis and reduces purine and pyrimidine synthesis resulting in reduction of cell prolifiration, and that this heme deficiency causes mitochondrial respiration defects and lethal oxidative stress (Vandekeere *et al.*, 2018). Others demonstrated that the mass balance within central carbon metabolism is disrupted upon inhibition of PHGDH independent of serine utilization (Reid *et al.*, 2018).

Discussion of the methodological approach

This study has successfully employed modern methods of genome modification, i.e., CRISPR/Cas9 and TALEN-based genome editing. While the CRISPR/Cas9 approach to directly introduce the point mutations of interest into the *PHGDH* gene with various gRNAs and repair templates was not successful, we were able to generate koHEK cells using this

technology. The commercial TALEN-based system contrasted with the CRISPR/Cas9 relativly easy. The *PHGDH* and its variants could be easily cloned into the donor vector. While the CRISPR/Cas9 nuclease activity is based on one enzyme producing the DSB, the TALEN nuclease is based on the FokI nuclease. FokI can only introduce a DSB when on both sides, sense, and an antisense strand, of the DNA a FokI has bound. In contrast to the CRISPR/Cas9, it is quite unlikely that both TALEN are binding at the same sequence, thereby reducing the risk of off-target effects (Nemudryi, Valetdinova, Medvedev, & Zakian, 2014; Zheng, Li, & Wang, 2020). The most important challenge, when using a TALEN approach, is the construction of good binding TALE domains to one specific target. This was not the case in this study since the goal was only to rescue the phenotype of the koHEK cells. In conclusion, according to the experience in this particular study, it could be recommended to use the CRISPR/Cas9 technology to introduce FS mutation and create thereby functional knock-out, but to introduce new genes, reporters, or to test variations of one gene the TALEN approach seemed to be the better choice.

A multimodal approach was used to characterize functionally *PHGDH* mutants. This allowed us to validate the findings and compare the value and validity of different methods. Following the generation of *PHGDH* knock-out HEK293 cells (koHEK), the study started with transient overexpression to study the effects of *PHGDH* mutations. However, it was noticed that enzymatic activity measurements in this system produced a quite wide variability and failed to discriminate exactly between different conditions of serine biosynthesis defects. Therefore, the stable transfection system was considered to achieve more homogeneous results. Indeed, the relatively high standard error observed in the current results from the transient transfection system, compared to the relatively lower standard error of the current results from the stable

transfection system is an indicator of the superiority of the stable transfection system (**Figure 18**, **Figure 19**, **Figure 22**, and **Figure 23**). In addition, it allowed us to study cultured transgenic lines in longer term; it must be taken into account. However, that the transient transfection is much easier and faster, whereas generating stable cell lines is tedious job and may not be practicable for quick functional validation of a variant.

In the beginning, these koHEK cells were transfected transiently by several constructs of pTagGFP-C-*PHGDH*. Considering that GFP is a big tag (about 27 kDa) one might wonder why GFP was selected as a tag for this work. The purification of the recombinant proteins was one of the rationales behind why the GFP-tag was used as μMACS™ GFP Isolation Kit was available at the time of the work and this purification system was established in the lab long before this work. The second reason was related to the initial idea of this work since the aim was to study the effect of special mutations on the phenotype of the cells or in other words if the difference in the phenotype results from the functional effects of these mutations. It is known that the right cellular localization of the recombinant proteins is the first indicator of the success of the protein expression system. For this reason, GFP as a fluorescent protein was selected in this study as an optimal tag to visualize the localization of the recombinant proteins, since GFP is easy to visualization with standard filter sets on a fluorescence microscope (Bhandari *et al.*, 2021; Fabrick & Hull, 2017).

Another point of discussion may be why HEK cells and not a prokaryotic expression system was used. Prokaryotic expression systems have the advantage of fast growth kinetics with a doubling time of about 20 min (Sezonov, Joseleau-Petit, & D'Ari, 2007). They are often chosen for generating recombinant proteins owing to their ease of use and relatively high yields. However, because of the biological tests and measurement of enzyme activity that were planned

in this study, the eukaryotic expression was needed. since for that goals, functional proteins and eukaryotic post-translational modifications and proper protein folding are required, which can be offered more precisely by the eukaryotic expression system. It is known that the processes of transcription and translation occur simultaneously in prokaryotes (Irastortza-Olaziregi & Amster-Choder, 2020). The translation of mRNA starts even before a mature mRNA transcript is fully synthesized. On the other hand, in eukaryotes, the transcription and the translation are spatially separated and occur sequentially, since the transcription happens in the nucleus and the translation occurs in the cytoplasm. After translation, the resulting polypeptides are modified in various ways to complete their structure, designate their location, or regulate their activity within the cell.

When the purification by µMACS™ GFP Isolation Kit was done by pH shift using triethylamine pH 11.8, then the purified recombinant proteins were non-functional in the enzyme assay presumably owing to the pH shift during the elution of the purified proteins from the µMACS column. For this reason, this approach was not further used, and non-purified proteins were used in this study for enzyme activity assay.

Limitations

The functional experiments performed in this study had several limitations: First of all, it was not possible in the given time to fully expand the stable transfection approach to the number of mutants tested in the transient expression system. Thereby, we are lacking a more detailed picture and a validation by a second method for the variants that were only tested in the transient expression system. However, it could be shown that both, the transient and the stable transfection systems are useful to study *PHGDH* mutants, while the multimodal approach increases the discriminatory power. The transient transfection system showed more variability

and appeared to overestimate the enzymatic activities of mutant PHGDH, probably related to the overexpression and the copy number of the plasmids per cell as was discussed before.

It also must be noted that only a limited number of *PHGDH* variants were tested. A more complete picture could probably have been reached, if more (e.g., all known) disease associated *PHGDH* variants were tested and compared. Currently, known disease-associated *PHGDH* variants published in the literature comprise 29 variants, of which only 9 were included in the transient expression experiment and 3 in the stable expression experiment. Additional pathogenic or likely pathogenic are listed in ClinVar, but no phenotype information is provided there. However, for studying the correlation between functional consequences of a mutant and phenotypic expression, the clear assignment of an individual variant to NLS or SBDNL is critical. This is, however, quite difficult, if only single patient exists with a certain variant, and particularly if a variant has only been found in a compound heterozygous state with a different change. We have attempted to make a rational assignment and classification of variants (**Appendix F Table 2**) (Abdelfattah *et al.*, 2020), but have to admit that some of these assignments are preliminary or ambiguous, and they may change with new cases reported.

It is likely that only quite small differences in residual enzyme activity may have quite strong impact on the phenotypic expression. Although it could be partially confirmed that NLS-associated variants have a lower level of functional capacity compared to variants associated with non-lethal phenotypes, it remains questionable, whether any single test can reach the power of discrimination to resolve these slight differences. It is also likely that there will not be a strict dichotomy between NLS and SBDNL but rather a spectrum of both, phenotypic severity, and related functional incapacity of the serine biosynthesis pathway.

Lastly, it cannot be excluded that other independent modifiers have an influence on phenotypic expression.

5.3 Conclusion and outlook

In summary, this study showed that pathogenic *PHGDH* and *PSAT1* variants lead to a spectrum of human disorders and that NLS per se has varying degrees of phenotypic expression likely representing the extreme end of a continuum. Seven novel mutations in *PSAT1* and two novel mutations in *PHGDH* were recognized in this study. In addition, the previously reported unsolved family with NLS has been solved in this study and the underlying genetic defect was detected. Therefore, this study supported no further the hypothesis of possible further genetic heterogeneity of NLS.

The study attempted for the first time to explore the functional effect of *PHGDH* variants related to NLS. A new multimodal *in vitro* approach was used to determine enzymatic activity and biological impact of *PHGDH* mutants, which is independent from the availability of patient cells. It was based on the successful generation of *PHGDH*-deficient HEK cells (koHEK) using targeting of the *PHGDH* locus by CRISPR/Cas9. These cells were used for transient overexpression of *PHGDH* mutants. In addition, targeting of the *AAVS1* locus in the generated koHEK cells using TALEN genome editing approach yielded robust CMV promoter-mediated stable transgene expression in koHEK cells.

Using these methods on a limited number of *PHGDH* variants, the pathogenic impact of newly and previously reported disease-associated variants could be reliably confirmed. In addition, this approach was partially successful in demonstrating a correlation between the individual residual enzyme activity of PHGDH mutants and the phenotypic expression of serine biosynthesis disorders. This was done through using two different transfection systems to determine the enzyme activity and the biological impact of some *PHGDH* variants. NLS-associated mutants tended to exhibit lower levels of enzymatic and biological activities

compared to SBDNL-associated variants, although not a strict dichotomy could be demonstrated.

This study identified limitations of the transient transfection system in the functional experiments, consistent with strongly varying results reported in the literature (**Table 7**). It also clarified the limitation of the used approach to discriminate exactly between different conditions of serine biosynthesis defects. Concerning *in vitro* disease modelling, the stable transfection cell line models turned out to be more reliable.

To reach a more complete picture of the functional spectrum of *PHGDH* variants and the correlation with disease expression, the stable transfection approach used in the work with the enzymatic and proliferation/viability assays could be expanded to a larger number of disease-associated variants. The toolset developed in this work could readily be used to provide experimental evidence for variants of uncertain significance, which may be of important significance for genetic counseling of affected families where variants of uncertain significance have been identified.

A similar approach could also be rolled out to *PSAT1* and *PSPH* and thereby complete the spectrum of SBDs from a scientific as we as from a clinical perspective.

6 References

Abdelfattah, F., Kariminejad, A., Kahlert, A. K., Morrison, P. J., Gumus, E., Mathews, K. D., . . . Schanze, D. (2020). Expanding the genotypic and phenotypic spectrum of severe serine biosynthesis disorders. *Hum Mutat, 41*(9), 1615-1628. doi:10.1002/humu.24067

Achouri, Y., Rider, M. H., Schaftingen, E. V., & Robbi, M. (1997). Cloning, sequencing and expression of rat liver 3-phosphoglycerate dehydrogenase. *Biochem J, 323 (Pt 2)*, 365-370. doi:10.1042/bj3230365

Acuna-Hidalgo, R., Schanze, D., Kariminejad, A., Nordgren, A., Kariminejad, M. H., Conner, P., . . . Zenker, M. (2014). Neu-Laxova syndrome is a heterogeneous metabolic disorder caused by defects in enzymes of the L-serine biosynthesis pathway. *Am J Hum Genet, 95*(3), 285-293. doi:10.1016/j.ajhg.2014.07.012

Agrotis, A., & Ketteler, R. (2015). A new age in functional genomics using CRISPR/Cas9 in arrayed library screening. *Front Genet, 6*, 300. doi:10.3389/fgene.2015.00300

Ali, A., Dhahouri, N. A., Almesmari, F. S. A., Fathalla, W. M., & Jasmi, F. A. (2021). Characterization of ETFDH and PHGDH Mutations in a Patient with Mild Glutaric Aciduria Type II and Serine Deficiency. *Genes, 12*(5), 703. Retrieved from **https://www.mdpi.com/2073-4425/12/5/703**

Arroyo-Olarte, R. D., Bravo Rodríguez, R., & Morales-Ríos, E. (2021). Genome Editing in Bacteria: CRISPR-Cas and Beyond. *Microorganisms, 9*(4). doi:10.3390/microorganisms9040844

Badakali, M., Badakali, A., & Dombale, V. (2012). Rare manifestations of Neu-Laxova syndrome. *Fetal Pediatr Pathol, 31*(1), 1-5. doi:10.3109/15513815.2011.618864

Baek, J. Y., Jun, D. Y., Taub, D., & Kim, Y. H. (2000). Assignment of human 3-phosphoglycerate dehydrogenase (PHGDH) to human chromosome band 1p12 by fluorescence in situ hybridization. *Cytogenet Cell Genet, 89*(1-2), 6-7. doi:10.1159/000015577

Baek, J. Y., Jun, D. Y., Taub, D., & Kim, Y. H. (2003). Characterization of human phosphoserine aminotransferase involved in the phosphorylated pathway of L-serine biosynthesis. *Biochem J, 373*(Pt 1), 191-200. doi:10.1042/bj20030144

Barrangou, R., Birmingham, A., Wiemann, S., Beijersbergen, R. L., Hornung, V., & Smith, A. (2015). Advances in CRISPR-Cas9 genome engineering: lessons learned from RNA interference. *Nucleic Acids Res, 43*(7), 3407-3419. doi:10.1093/nar/gkv226

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., . . . Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science, 315*(5819), 1709-1712. doi:10.1126/science.1138140

Bassett, A., & Liu, J. L. (2014). CRISPR/Cas9 mediated genome engineering in Drosophila. *Methods, 69*(2), 128-136. doi:10.1016/j.ymeth.2014.02.019

Becker, S., & Boch, J. (2021). TALE and TALEN genome editing technologies. *Gene and Genome Editing, 2*, 100007. doi:**https://doi.org/10.1016/j.ggedit.2021.100007**

Benke, P. J., Hidalgo, R. J., Braffman, B. H., Jans, J., Gassen, K., Sunbul, R., & El-Hattab, A. W. (2017). Infantile Serine Biosynthesis Defect Due to Phosphoglycerate Dehydrogenase Deficiency: Variability in Phenotype and Treatment Response, Novel Mutations, and Diagnostic Challenges. *J Child Neurol, 32*(6), 543-549. doi:10.1177/0883073817690094

Bhandari, B. K., Lim, C. S., Remus, D. M., Chen, A., van Dolleweerd, C., & Gardner, P. P. (2021). Analysis of 11,430 recombinant protein production experiments reveals that protein yield is tunable by synonymous codon changes of translation initiation sites. *PLoS Comput Biol, 17*(10), e1009461. doi:10.1371/journal.pcbi.1009461

Bibikova, M., Beumer, K., Trautman, J. K., & Carroll, D. (2003). Enhancing gene targeting with designed zinc finger nucleases. *Science, 300*(5620), 764. doi:10.1126/science.1079512

Bibikova, M., Carroll, D., Segal, D. J., Trautman, J. K., Smith, J., Kim, Y. G., & Chandrasegaran, S. (2001). Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol Cell Biol, 21*(1), 289-297. doi:10.1128/mcb.21.1.289-297.2001

Boch, J., & Bonas, U. (2010). Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annu Rev Phytopathol, 48*, 419-436. doi:10.1146/annurev-phyto-080508-081936

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., . . . Bonas, U. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. *Science, 326*(5959), 1509-1512. doi:10.1126/science.1178811

Bogdanove, A. J., & Voytas, D. F. (2011). TAL effectors: customizable proteins for DNA targeting. *Science, 333*(6051), 1843-1846. doi:10.1126/science.1204094

Bolotin, A., Quinquis, B., Sorokin, A., & Ehrlich, S. D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology (Reading), 151*(Pt 8), 2551-2561. doi:10.1099/mic.0.28048-0

Bourque, D. K., Cloutier, M., Kernohan, K. D., Bareke, E., Grynspan, D., Michaud, J., . . . Boycott, K. M. (2019). Neu-Laxova syndrome presenting prenatally with increased nuchal translucency and cystic hygroma: The utility of exome sequencing in deciphering the diagnosis. *Am J Med Genet A, 179*(5), 813-816. doi:10.1002/ajmg.a.61076

Brassier, A., Valayannopoulos, V., Bahi-Buisson, N., Wiame, E., Hubert, L., Boddaert, N., . . . de Lonlay, P. (2016). Two new cases of serine deficiency disorders treated with l-serine. *Eur J Paediatr Neurol, 20*(1), 53-60. doi:10.1016/j.ejpn.2015.10.007

Brouns, S. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J., Snijders, A. P., . . . van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science, 321*(5891), 960-964. doi:10.1126/science.1159689

Bultmann, S., Morbitzer, R., Schmidt, C. S., Thanisch, K., Spada, F., Elsaesser, J., . . . Leonhardt, H. (2012). Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. *Nucleic Acids Res, 40*(12), 5368-5377. doi:10.1093/nar/gks199

Byers, H. M., Bennett, R. L., Malouf, E. A., Weiss, M. D., Feng, J., Scott, C. R., & Jayadev, S. (2016). Novel Report of Phosphoserine Phosphatase Deficiency in an Adult with Myeloneuropathy and Limb Contractures. *JIMD Rep, 30*, 103-108. doi:10.1007/8904_2015_510

Capecchi, M. (1990). Gene targeting. How efficient can you get? *Nature, 348*(6297), 109. doi:10.1038/348109a0

Capecchi, M. R. (1989). Altering the genome by homologous recombination. *Science, 244*(4910), 1288-1292. doi:10.1126/science.2660260

Carder, K. R., Fitzpatrick, J. E., & Weston, W. L. (2003). What syndrome is this? Neu-Laxova syndrome. *Pediatr Dermatol, 20*(1), 78-80. doi:10.1046/j.1525-1470.2003.03017.x

Cavole, T. R., Perrone, E., Lucena de Castro, F. S. C., Alvarez Perez, A. B., Waitzberg, A. F. L., & Cernach, M. (2020). Clinical, molecular, and pathological findings in a Neu-Laxova syndrome stillborn: A Brazilian case report. *Am J Med Genet A*. doi:10.1002/ajmg.a.61559

Cermak, T., Doyle, E. L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., . . . Voytas, D. F. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res, 39*(12), e82. doi:10.1093/nar/gkr218

Cho, H. M., Jun, D. Y., Bae, M. A., Ahn, J. D., & Kim, Y. H. (2000). Nucleotide sequence and differential expression of the human 3-phosphoglycerate dehydrogenase gene. *Gene, 245*(1), 193-201. doi:10.1016/s0378-1119(00)00009-3

Cho, S. W., Kim, S., Kim, Y., Kweon, J., Kim, H. S., Bae, S., & Kim, J. S. (2014). Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res, 24*(1), 132-141. doi:10.1101/gr.162339.113

Choulika, A., Perrin, A., Dujon, B., & Nicolas, J. F. (1995). Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of Saccharomyces cerevisiae. *Mol Cell Biol, 15*(4), 1968-1973. doi:10.1128/mcb.15.4.1968

Christian, M., Cermak, T., Doyle, E. L., Schmidt, C., Zhang, F., Hummel, A., . . . Voytas, D. F. (2010). Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics, 186*(2), 757-761. doi:10.1534/genetics.110.120717

Collet, J. F., Gerin, I., Rider, M. H., Veiga-da-Cunha, M., & Van Schaftingen, E. (1997). Human L-3-phosphoserine phosphatase: sequence, expression and evidence for a phosphoenzyme intermediate. *FEBS Lett, 408*(3), 281-284. doi:10.1016/s0014-5793(97)00438-9

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., . . . Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science, 339*(6121), 819-823. doi:10.1126/science.1231143

Coşkun, T., Aydin, H. I., Kiliç, M., Dursun, A., Haliloğlu, G., Topaloğlu, H., . . . de Koning, T. J. (2009). 3-phosphoglycerate dehydrogenase deficiency: a case report of a treatable cause of seizures. *Turk J Pediatr, 51*(6), 587-592.

Coto-Puckett, W. L., Gilbert-Barness, E., Steelman, C. K., Stuart, T., Robinson, H. B., & Shehata, B. M. (2010). A spectrum of phenotypical expression OF Neu-Laxova syndrome: Three case reports and a review of the literature. *Fetal Pediatr Pathol, 29*(2), 108-119. doi:10.3109/15513811003620914

Curry, C. J. (1982). Further comments on the Neu-Laxova syndrome. *Am J Med Genet, 13*(4), 441-444. doi:10.1002/ajmg.1320130414

de Koning, T. J., Duran, M., Dorland, L., Gooskens, R., Van Schaftingen, E., Jaeken, J., . . . Poll-The, B. T. (1998). Beneficial effects of L-serine and glycine in the management of seizures in 3-phosphoglycerate dehydrogenase deficiency. *Ann Neurol, 44*(2), 261-265. doi:10.1002/ana.410440219

DeBerardinis, R. J. (2011). Serine metabolism: some tumors take the road less traveled. *Cell Metab, 14*(3), 285-286. doi:10.1016/j.cmet.2011.08.004

Debs, S., Ferreira, C. R., Groden, C., Kim, H. J., King, K. A., King, M. C., . . . Soldatos, A. (2021). Adult diagnosis of congenital serine biosynthesis defect: A treatable cause of progressive neuropathy. *Am J Med Genet A, 185*(7), 2102-2107. doi:10.1002/ajmg.a.62245

DeKelver, R. C., Choi, V. M., Moehle, E. A., Paschon, D. E., Hockemeyer, D., Meijsing, S. H., . . . Urnov, F. D. (2010). Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. *Genome Res, 20*(8), 1133-1142. doi:10.1101/gr.106773.110

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., . . . Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature, 471*(7340), 602-607. doi:10.1038/nature09886

Dow, L. E. (2015). Modeling Disease In Vivo With CRISPR/Cas9. *Trends Mol Med, 21*(10), 609-621. doi:10.1016/j.molmed.2015.07.006

Eade, K., Gantner, M. L., Hostyk, J. A., Nagasaki, T., Giles, S., Fallon, R., . . . Allikmets, R. (2021). Serine biosynthesis defect due to haploinsufficiency of PHGDH causes retinal disease. *Nat Metab, 3*(3), 366-377. doi:10.1038/s42255-021-00361-3

Ejeckam, G. G., Wadhwa, J. K., Williams, J. P., & Lacson, A. G. (1986). Neu-Laxova syndrome: report of two cases. *Pediatr Pathol, 5*(3-4), 295-306. Retrieved from **https://www.ncbi.nlm.nih.gov/pubmed/3786261**

El-Hattab, A. W., Shaheen, R., Hertecant, J., Galadari, H. I., Albaqawi, B. S., Nabil, A., & Alkuraya, F. S. (2016). On the phenotypic spectrum of serine biosynthesis defects. *J Inherit Metab Dis, 39*(3), 373-381. doi:10.1007/s10545-016-9921-5

Fabrick, J. A., & Hull, J. J. (2017). Transient Expression and Cellular Localization of Recombinant Proteins in Cultured Insect Cells. *J Vis Exp*(122). doi:10.3791/55756

Fitch, N., Resch, L., & Rochon, L. (1982). The Neu-Laxova syndrome: comments on syndrome identification. *Am J Med Genet, 13*(4), 445-452. doi:10.1002/ajmg.1320130415

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., & Sander, J. D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol, 31*(9), 822-826. doi:10.1038/nbt.2623

Furuya, S. (2008). An essential role for de novo biosynthesis of L-serine in CNS development. *Asia Pac J Clin Nutr, 17 Suppl 1*, 312-315. Retrieved from **https://www.ncbi.nlm.nih.gov/pubmed/18296366**

Furuya, S., Yoshida, K., Kawakami, Y., Yang, J. H., Sayano, T., Azuma, N., . . . Hirabayashi, Y. (2008). Inactivation of the 3-phosphoglycerate dehydrogenase gene in mice: changes in gene expression and associated regulatory networks resulting from serine deficiency. *Funct Integr Genomics, 8*(3), 235-249. doi:10.1007/s10142-007-0072-5

Glinton, K. E., Benke, P. J., Lines, M. A., Geraghty, M. T., Chakraborty, P., Al-Dirbashi, O. Y., . . . El-Hattab, A. W. (2018). Disturbed phospholipid metabolism in serine biosynthesis defects revealed by metabolomic profiling. *Mol Genet Metab, 123*(3), 309-316. doi:10.1016/j.ymgme.2017.12.009

Grant, G. A. (2018). D-3-Phosphoglycerate Dehydrogenase. *Frontiers in Molecular Biosciences, 5*. doi:10.3389/fmolb.2018.00110

Grant, G. A. (2018). D-3-Phosphoglycerate Dehydrogenase. *Front Mol Biosci, 5*, 110. doi:10.3389/fmolb.2018.00110

Guex, N., & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis, 18*(15), 2714-2723. doi:10.1002/elps.1150181505

Hart, C. E., Race, V., Achouri, Y., Wiame, E., Sharrard, M., Olpin, S. E., . . . Van Schaftingen, E. (2007). Phosphoserine aminotransferase deficiency: a novel disorder of the serine biosynthesis pathway. *Am J Hum Genet, 80*(5), 931-937. doi:10.1086/517888

Hausler, M. G., Jaeken, J., Monch, E., & Ramaekers, V. T. (2001). Phenotypic heterogeneity and adverse effects of serine treatment in 3-phosphoglycerate dehydrogenase deficiency: report on two siblings. *Neuropediatrics, 32*(4), 191-195. doi:10.1055/s-2001-17373

Hester, G., Stark, W., Moser, M., Kallen, J., Markovic-Housley, Z., & Jansonius, J. N. (1999). Crystal structure of phosphoserine aminotransferase from Escherichia coli at 2.3 A resolution: comparison of the unligated enzyme and a complex with alpha-methyl-l-glutamate. *J Mol Biol, 286*(3), 829-850. doi:10.1006/jmbi.1998.2506

Hockemeyer, D., Soldner, F., Cook, E. G., Gao, Q., Mitalipova, M., & Jaenisch, R. (2008). A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. *Cell Stem Cell, 3*(3), 346-353. doi:10.1016/j.stem.2008.08.014

Horn, D., Muller, D., Thiele, H., & Kunze, J. (1997). Extreme microcephaly, severe growth and mental retardation, flexion contractures, and ichthyotic skin in two brothers: a new syndrome or mild form of Neu-Laxova syndrome? *Clin Dysmorphol, 6*(4), 323-328. Retrieved from **https://www.ncbi.nlm.nih.gov/pubmed/9354840**

Hryhorowicz, M., Lipiński, D., Zeyland, J., & Słomski, R. (2017). CRISPR/Cas9 Immune System as a Tool for Genome Engineering. *Archivum immunologiae et therapiae experimentalis, 65*(3), 233-240. doi:10.1007/s00005-016-0427-5

Hsu, P. D., Lander, E. S., & Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell, 157*(6), 1262-1278. doi:10.1016/j.cell.2014.05.010

Irastortza-Olaziregi, M., & Amster-Choder, O. (2020). Coupled Transcription-Translation in Prokaryotes: An Old Couple With New Surprises. *Front Microbiol, 11*, 624830. doi:10.3389/fmicb.2020.624830

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., & Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *J Bacteriol, 169*(12), 5429-5433. doi:10.1128/jb.169.12.5429-5433.1987

Jaeken, J., Detheux, M., Fryns, J. P., Collet, J. F., Alliet, P., & Van Schaftingen, E. (1997). Phosphoserine phosphatase deficiency in a patient with Williams syndrome. *J Med Genet, 34*(7), 594-596. doi:10.1136/jmg.34.7.594

Jaeken, J., Detheux, M., Van Maldergem, L., Foulon, M., Carchon, H., & Van Schaftingen, E. (1996). 3-Phosphoglycerate dehydrogenase deficiency: an inborn error of serine biosynthesis. *Arch Dis Child, 74*(6), 542-545. doi:10.1136/adc.74.6.542

Jansen, R., Embden, J. D., Gaastra, W., & Schouls, L. M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol, 43*(6), 1565-1575. doi:10.1046/j.1365-2958.2002.02839.x

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

Jo, Y. I., Suresh, B., Kim, H., & Ramakrishna, S. (2015). CRISPR/Cas9 system as an innovative genetic engineering tool: Enhancements in sequence specificity and delivery methods. *Biochim Biophys Acta, 1856*(2), 234-243. doi:10.1016/j.bbcan.2015.09.003

John, R. A. (1995). Pyridoxal phosphate-dependent enzymes. *Biochim Biophys Acta, 1248*(2), 81-96. doi:10.1016/0167-4838(95)00025-p

Joung, J. K., & Sander, J. D. (2013). TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol, 14*(1), 49-55. doi:10.1038/nrm3486

Kalhan, S. C., & Hanson, R. W. (2012). Resurgence of Serine: An Often Neglected but Indispensable Amino Acid*. *Journal of Biological Chemistry, 287*(24), 19786-19791. doi:**https://doi.org/10.1074/jbc.R112.357194**

Kapoor, R., Thakur, S., Kapoor, A., Kapoor, S., Kalra, A., & Kapoor, A. (2021). Neu–Laxova's Syndrome: A Case Report of a Fetus with Novel Mutation in PHGDH Gene and a Literature Review. *J Pediatr Genet*(EFirst).

Karimi-Nejad, M. H., Khajavi, H., Gharavi, M. J., & Karimi-Nejad, R. (1987). Neu-Laxova syndrome: report of a case and comments. *Am J Med Genet, 28*(1), 17-23. doi:10.1002/ajmg.1320280104

Kim, H., & Kim, J.-S. (2014). A guide to genome engineering with programmable nucleases. *Nature Reviews Genetics, 15*(5), 321-334. doi:10.1038/nrg3686

Kim, Y. G., Cha, J., & Chandrasegaran, S. (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences of the United States of America, 93*(3), 1156-1160. doi:10.1073/pnas.93.3.1156

Klomp, L. W., de Koning, T. J., Malingre, H. E., van Beurden, E. A., Brink, M., Opdam, F. L., . . . Berger, R. (2000). Molecular characterization of 3-phosphoglycerate dehydrogenase deficiency--a neurometabolic disorder associated with reduced L-serine biosynthesis. *Am J Hum Genet, 67*(6), 1389-1399. doi:10.1086/316886

Kotin, R. M., Linden, R. M., & Berns, K. I. (1992). Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination. *Embo j, 11*(13), 5071-5078. doi:10.1002/j.1460-2075.1992.tb05614.x

Kraoua, I., Wiame, E., Kraoua, L., Nasrallah, F., Benrhouma, H., Rouissi, A., . . . Gouider-Khouja, N. (2013). 3-Phosphoglycerate dehydrogenase deficiency: description of two new cases in Tunisia and review of the literature. *Neuropediatrics, 44*(5), 281-285. doi:10.1055/s-0033-1338133

Labun, K., Montague, T. G., Gagnon, J. A., Thyme, S. B., & Valen, E. (2016). CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res, 44*(W1), W272-276. doi:10.1093/nar/gkw398

Laxova, R., Ohara, P. T., & Timothy, J. A. (1972). A further example of a lethal autosomal recessive condition in sibs. *J Ment Defic Res, 16*(2), 139-143. doi:10.1111/j.1365-2788.1972.tb01585.x

Li, H., Yang, Y., Hong, W., Huang, M., Wu, M., & Zhao, X. (2020). Applications of genome editing technology in the targeted therapy of human diseases: mechanisms, advances and prospects. *Signal Transduction and Targeted Therapy, 5*(1), 1. doi:10.1038/s41392-019-0089-y

Locasale, J. W., Grassian, A. R., Melman, T., Lyssiotis, C. A., Mattaini, K. R., Bass, A. J., . . . Vander Heiden, M. G. (2011). Phosphoglycerate dehydrogenase diverts glycolytic flux and contributes to oncogenesis. *Nat Genet, 43*(9), 869-874. doi:10.1038/ng.890

Lombardo, A., Cesana, D., Genovese, P., Di Stefano, B., Provasi, E., Colombo, D. F., . . . Naldini, L. (2011). Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nat Methods, 8*(10), 861-869. doi:10.1038/nmeth.1674

Mak, A. N., Bradley, P., Cernadas, R. A., Bogdanove, A. J., & Stoddard, B. L. (2012). The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science, 335*(6069), 716-719. doi:10.1126/science.1216211

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., . . . Koonin, E. V. (2011). Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol, 9*(6), 467-477. doi:10.1038/nrmicro2577

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., . . . Church, G. M. (2013). RNA-guided human genome engineering via Cas9. *Science, 339*(6121), 823-826. doi:10.1126/science.1232033

Manning, M. A., Cunniff, C. M., Colby, C. E., El-Sayed, Y. Y., & Hoyme, H. E. (2004). Neu-Laxova syndrome: detailed prenatal diagnostic and post-mortem findings and literature review. *Am J Med Genet A, 125A*(3), 240-249. doi:10.1002/ajmg.a.20467

Marraffini, L. A., & Sontheimer, E. J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science, 322*(5909), 1843-1845. doi:10.1126/science.1165771

Mattaini, K. R., Sullivan, M. R., & Vander Heiden, M. G. (2016). The importance of serine metabolism in cancer. *J Cell Biol, 214*(3), 249-257. doi:10.1083/jcb.201604085

Mattos, E. P., Silva, A. A., Magalhaes, J. A., Leite, J. C., Leistner-Segal, S., Gus-Kessler, R., . . . Sanseverino, M. T. (2015). Identification of a premature stop codon mutation in the PHGDH gene in severe Neu-Laxova syndrome-evidence for phenotypic variability. *Am J Med Genet A, 167*(6), 1323-1329. doi:10.1002/ajmg.a.36930

Meneret, A., Wiame, E., Marelli, C., Lenglet, T., Van Schaftingen, E., & Sedel, F. (2012). A serine synthesis defect presenting with a Charcot-Marie-Tooth-like polyneuropathy. *Arch Neurol, 69*(7), 908-911. doi:10.1001/archneurol.2011.1526

Mishra, V., Ali, V., Nozaki, T., & Bhakuni, V. (2010). Entamoeba histolytica Phosphoserine aminotransferase (EhPSAT): insights into the structure-function relationship. *BMC Res Notes, 3*, 52. doi:10.1186/1756-0500-3-52

Mishra, V., Kumar, A., Ali, V., Nozaki, T., Zhang, K. Y., & Bhakuni, V. (2012). Glu-108 is essential for subunit assembly and dimer stability of D-phosphoglycerate dehydrogenase from Entamoeba histolytica. *Mol Biochem Parasitol, 181*(2), 117-124. doi:10.1016/j.molbiopara.2011.10.008

Mojica, F. J., Díez-Villaseñor, C., García-Martínez, J., & Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol, 60*(2), 174-182. doi:10.1007/s00239-004-0046-3

Mojica, F. J., Díez-Villaseñor, C., Soria, E., & Juez, G. (2000). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol Microbiol, 36*(1), 244-246. doi:10.1046/j.1365-2958.2000.01838.x

Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., & Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology (Reading), 155*(Pt 3), 733-740. doi:10.1099/mic.0.023960-0

Monies, D., Abouelhoda, M., AlSayed, M., Alhassnan, Z., Alotaibi, M., Kayyali, H., . . . Alkuraya, F. S. (2017). The landscape of genetic diseases in Saudi Arabia based on the first 1000 diagnostic panels and exomes. *Hum Genet, 136*(8), 921-939. doi:10.1007/s00439-017-1821-8

Monies, D., Abouelhoda, M., Assoum, M., Moghrabi, N., Rafiullah, R., Almontashiri, N., . . . Alkuraya, F. S. (2019). Lessons Learned from Large-Scale, First-Tier Clinical Exome Sequencing in a Highly Consanguineous Population. *Am J Hum Genet, 104*(6), 1182-1201. doi:10.1016/j.ajhg.2019.04.011

Moscou, M. J., & Bogdanove, A. J. (2009). A simple cipher governs DNA recognition by TAL effectors. *Science, 326*(5959), 1501. doi:10.1126/science.1178817

Murtas, G., Marcone, G. L., Sacchi, S., & Pollegioni, L. (2020). L-serine synthesis via the phosphorylated pathway in humans. *Cell Mol Life Sci, 77*(24), 5131-5148. doi:10.1007/s00018-020-03574-z

Naito, Y., Hino, K., Bono, H., & Ui-Tei, K. (2015). CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics, 31*(7), 1120-1123. doi:10.1093/bioinformatics/btu743

Naveed, Manjunath, C. S., & Sreenivas, V. (1990). New manifestations of Neu-Laxova syndrome. *Am J Med Genet, 35*(1), 55-59. doi:10.1002/ajmg.1320350110

Nemudryi, A. A., Valetdinova, K. R., Medvedev, S. P., & Zakian, S. M. (2014). TALEN and CRISPR/Cas Genome Editing Systems: Tools of Discovery. *Acta Naturae, 6*(3), 19-40.

Neu, R. L., Kajii, T., Gardner, L. I., & Nagyfy, S. F. (1971). A lethal syndrome of microcephaly with multiple congenital anomalies in three siblings. *Pediatrics, 47*(3), 610-612. Retrieved from **https://www.ncbi.nlm.nih.gov/pubmed/5547878**

Ni, C., Cheng, R. H., Zhang, J., Liang, J. Y., Wei, R. Q., Li, M., & Yao, Z. R. (2019). Novel and recurrent PHGDH and PSAT1 mutations in Chinese patients with Neu-Laxova syndrome. *Eur J Dermatol, 29*(6), 641-646. doi:10.1684/ejd.2019.3673

Ostrovskaya, T. I., & Lazjuk, G. I. (1988). Cerebral abnormalities in the Neu-Laxova syndrome. *Am J Med Genet, 30*(3), 747-756. doi:10.1002/ajmg.1320300308

Ou, Y., Wang, S. J., Jiang, L., Zheng, B., & Gu, W. (2015). p53 Protein-mediated regulation of phosphoglycerate dehydrogenase (PHGDH) is crucial for the apoptotic response upon serine starvation. *J Biol Chem, 290*(1), 457-466. doi:10.1074/jbc.M114.616359

Pind, S., Slominski, E., Mauthe, J., Pearlman, K., Swoboda, K. J., Wilkins, J. A., . . . Natowicz, M. R. (2002). V490M, a common mutation in 3-phosphoglycerate dehydrogenase deficiency, causes enzyme deficiency by decreasing the yield of mature enzyme. *J Biol Chem, 277*(9), 7136-7143. doi:10.1074/jbc.M111419200

Pineda, M., Vilaseca, M. A., Artuch, R., Santos, S., Garcia Gonzalez, M. M., Aracil, A., . . . Jaeken, J. (2000). 3-phosphoglycerate dehydrogenase deficiency in a patient with West syndrome. *Dev Med Child Neurol, 42*(9), 629-633. doi:10.1017/s0012162200001171

Poli, A., Vial, Y., Haye, D., Passemard, S., Schiff, M., Nasser, H., . . . Verloes, A. (2017). Phosphoglycerate dehydrogenase (PHGDH) deficiency without epilepsy mimicking primary microcephaly. *Am J Med Genet A, 173*(7), 1936-1942. doi:10.1002/ajmg.a.38217

Pollari, S., Käkönen, S. M., Edgren, H., Wolf, M., Kohonen, P., Sara, H., . . . Kallioniemi, O. (2011). Enhanced serine production by bone metastatic breast cancer cells stimulates osteoclastogenesis. *Breast Cancer Res Treat, 125*(2), 421-430. doi:10.1007/s10549-010-0848-5

Possemato, R., Marks, K. M., Shaul, Y. D., Pacold, M. E., Kim, D., Birsoy, K., . . . Sabatini, D. M. (2011). Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature, 476*(7360), 346-350. doi:10.1038/nature10350

Prelich, G. (2012). Gene overexpression: uses, mechanisms, and interpretation. *Genetics, 190*(3), 841-854. doi:10.1534/genetics.111.136911

Ran, F. A., Hsu, P. D., Lin, C. Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., . . . Zhang, F. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell, 154*(6), 1380-1389. doi:10.1016/j.cell.2013.08.021

Ravez, S., Spillier, Q., Marteau, R., Feron, O., & Frédérick, R. (2017). Challenges and Opportunities in the Development of Serine Synthetic Pathway Inhibitors for Cancer Therapy. *J Med Chem, 60*(4), 1227-1237. doi:10.1021/acs.jmedchem.6b01167

Reid, M. A., Allen, A. E., Liu, S., Liberti, M. V., Liu, P., Liu, X., . . . Locasale, J. W. (2018). Serine synthesis through PHGDH coordinates nucleotide levels by maintaining central carbon metabolism. *Nat Commun, 9*(1), 5442. doi:10.1038/s41467-018-07868-6

Reyon, D., Tsai, S. Q., Khayter, C., Foden, J. A., Sander, J. D., & Joung, J. K. (2012). FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol, 30*(5), 460-465. doi:10.1038/nbt.2170

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., . . . Committee, A. L. Q. A. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med, 17*(5), 405-424. doi:10.1038/gim.2015.30

Rouet, P., Smih, F., & Jasin, M. (1994). Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol Cell Biol, 14*(12), 8096-8106. doi:10.1128/mcb.14.12.8096-8106.1994

Rouzbahani, L. (1995). New manifestations in an infant with Neu Laxova syndrome. *Am J Med Genet, 56*(2), 239-240. doi:10.1002/ajmg.1320560225

Sadelain, M., Papapetrou, E. P., & Bushman, F. D. (2011). Safe harbours for the integration of new DNA in the human genome. *Nat Rev Cancer, 12*(1), 51-58. doi:10.1038/nrc3179

Sayle, R. A., & Milner-White, E. J. (1995). RASMOL: biomolecular graphics for all. *Trends Biochem Sci, 20*(9), 374. doi:10.1016/s0968-0004(00)89080-5

Scherer, S., & Davis, R. W. (1979). Replacement of chromosome segments with altered DNA sequences constructed in vitro. *Proceedings of the National Academy of Sciences of the United States of America, 76*(10), 4951-4955. doi:10.1073/pnas.76.10.4951

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat Methods, 9*(7), 671-675. doi:10.1038/nmeth.2089

Seruggia, D., & Montoliu, L. (2014). The new CRISPR-Cas system: RNA-guided genome engineering to efficiently produce any desired genetic alteration in animals. *Transgenic Res, 23*(5), 707-716. doi:10.1007/s11248-014-9823-y

Sezonov, G., Joseleau-Petit, D., & D'Ari, R. (2007). Escherichia coli physiology in Luria-Bertani broth. *J Bacteriol, 189*(23), 8746-8749. doi:10.1128/jb.01368-07

Shaheen, R., Rahbeeni, Z., Alhashem, A., Faqeih, E., Zhao, Q., Xiong, Y., . . . Alkuraya, F. S. (2014). Neu-Laxova syndrome, an inborn error of serine metabolism, is caused by mutations in PHGDH. *Am J Hum Genet, 94*(6), 898-904. doi:10.1016/j.ajhg.2014.04.015

Shen, Y., Peng, Y., Huang, P., Zheng, Y., Li, S., Jiang, K., . . . Hong, D. (2022). Juvenile-onset PSAT1-related neuropathy: A milder phenotype of serine deficiency disorder. *Front Genet, 13*. doi:10.3389/fgene.2022.949038

Shivarajan, M. A., Suresh, S., Jagadeesh, S., Lata, S., & Bhat, L. (2003). Second trimester diagnosis of Neu Laxova syndrome. *Prenat Diagn, 23*(1), 21-24. doi:10.1002/pd.485

Shrivastav, M., De Haro, L. P., & Nickoloff, J. A. (2008). Regulation of DNA double-strand break repair pathway choice. *Cell Res, 18*(1), 134-147. doi:10.1038/cr.2007.111

Sirr, A., Lo, R. S., Cromie, G. A., Scott, A. C., Ashmead, J., Heyesus, M., & Dudley, A. M. (2020). A yeast-based complementation assay elucidates the functional impact of 200 missense variants in human PSAT1. *J Inherit Metab Dis*. doi:10.1002/jimd.12227

Smith, C., Gore, A., Yan, W., Abalde-Atristain, L., Li, Z., He, C., . . . Ye, Z. (2014). Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs. *Cell Stem Cell, 15*(1), 12-13. doi:10.1016/j.stem.2014.06.011

Smith, J. R., Maguire, S., Davis, L. A., Alexander, M., Yang, F., Chandran, S., . . . Pedersen, R. A. (2008). Robust, persistent transgene expression in human embryonic stem cells is achieved with AAVS1-targeted integration. *Stem Cells, 26*(2), 496-504. doi:10.1634/stemcells.2007-0039

Snell, K. (1986). The duality of pathways for serine biosynthesis is a fallacy. *Trends in Biochemical Sciences, 11*, 241-243.

Snell, K., & Weber, G. (1986). Enzymic imbalance in serine metabolism in rat hepatomas. *Biochem J, 233*(2), 617-620. doi:10.1042/bj2330617

Surks, H. K., Richards, C. T., & Mendelsohn, M. E. (2003). Myosin phosphatase-Rho interacting protein. A new member of the myosin phosphatase complex that directly binds RhoA. *J Biol Chem, 278*(51), 51484-51493. doi:10.1074/jbc.M305622200

Tabatabaie, L., de Koning, T. J., Geboers, A. J., van den Berg, I. E., Berger, R., & Klomp, L. W. (2009). Novel mutations in 3-phosphoglycerate dehydrogenase (PHGDH) are distributed throughout the protein and result in altered enzyme kinetics. *Hum Mutat, 30*(5), 749-756. doi:10.1002/humu.20934

Tabatabaie, L., Klomp, L. W., Berger, R., & de Koning, T. J. (2010). L-serine synthesis in the central nervous system: a review on serine deficiency disorders. *Mol Genet Metab, 99*(3), 256-262. doi:10.1016/j.ymgme.2009.10.012

Tabatabaie, L., Klomp, L. W., Rubio-Gozalbo, M. E., Spaapen, L. J., Haagen, A. A., Dorland, L., & de Koning, T. J. (2011). Expanding the clinical spectrum of 3-phosphoglycerate dehydrogenase deficiency. *J Inherit Metab Dis, 34*(1), 181-184. doi:10.1007/s10545-010-9249-5

Takeichi, T., Okuno, Y., Kawamoto, A., Inoue, T., Nagamoto, E., Murase, C., . . . Akiyama, M. (2018). Reduction of stratum corneum ceramides in Neu-Laxova syndrome caused by phosphoglycerate dehydrogenase deficiency. *J Lipid Res, 59*(12), 2413-2420. doi:10.1194/jlr.P087536

Torres-Ruiz, R., & Rodriguez-Perales, S. (2017). CRISPR-Cas9 technology: applications and human disease modelling. *Brief Funct Genomics, 16*(1), 4-12. doi:10.1093/bfgp/elw025

Ugras, M., Kocak, G., & Ozcan, H. (2006). Neu-Laxova syndrome: a case report and review of the literature. *J Eur Acad Dermatol Venereol, 20*(9), 1126-1128. doi:10.1111/j.1468-3083.2006.01645.x

Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S., & Gregory, P. D. (2010). Genome editing with engineered zinc finger nucleases. *Nature Reviews Genetics, 11*(9), 636-646. doi:10.1038/nrg2842

van der Crabben, S. N., Verhoeven-Duif, N. M., Brilstra, E. H., Van Maldergem, L., Coskun, T., Rubio-Gozalbo, E., . . . de Koning, T. J. (2013). An update on serine deficiency disorders. *J Inherit Metab Dis, 36*(4), 613-619. doi:10.1007/s10545-013-9592-4

van der Zel, A., Lam, H. M., & Winkler, M. E. (1989). Extensive homology between the Escherichia coli K-12 SerC(PdxF) aminotransferase and a protein encoded by a progesterone-induced mRNA in rabbit and human endometria. *Nucleic Acids Res, 17*(20), 8379. doi:10.1093/nar/17.20.8379

Vandekeere, S., Dubois, C., Kalucka, J., Sullivan, M. R., García-Caballero, M., Goveia, J., . . . Carmeliet, P. (2018). Serine Synthesis via PHGDH Is Essential for Heme Production in Endothelial Cells. *Cell Metab, 28*(4), 573-587.e513. doi:10.1016/j.cmet.2018.06.009

Veeranna, & Shetty, K. T. (1990). Phosphoserine phosphatase of human brain: partial purification, characterization, regional distribution, and effect of certain modulators including psychoactive drugs. *Neurochem Res, 15*(12), 1203-1210. doi:10.1007/bf01208581

Veiga-da-Cunha, M., Collet, J. F., Prieur, B., Jaeken, J., Peeraer, Y., Rabbijns, A., & Van Schaftingen, E. (2004). Mutations responsible for 3-phosphoserine phosphatase deficiency. *Eur J Hum Genet, 12*(2), 163-166. doi:10.1038/sj.ejhg.5201083

Vincent, J. B., Jamil, T., Rafiq, M. A., Anwar, Z., Ayaz, M., Hameed, A., . . . Ayub, M. (2015). Phosphoserine phosphatase (PSPH) gene mutation in an intellectual disability family from Pakistan. *Clin Genet, 87*(3), 296-298. doi:10.1111/cge.12445

Wang, W., Wu, Z., Dai, Z., Yang, Y., Wang, J., & Wu, G. (2013). Glycine metabolism in animals and humans: implications for nutrition and health. *Amino Acids, 45*(3), 463-477. doi:10.1007/s00726-013-1493-1

Xu, H., Qing, X., Wang, Q., Li, C., & Lai, L. (2021). Dimerization of PHGDH via the catalytic unit is essential for its enzymatic function. *J Biol Chem, 296*, 100572. doi:10.1016/j.jbc.2021.100572

Yang, J. H., Wada, A., Yoshida, K., Miyoshi, Y., Sayano, T., Esaki, K., . . . Furuya, S. (2010). Brain-specific Phgdh deletion reveals a pivotal role for L-serine biosynthesis in controlling the level of D-serine, an N-methyl-D-aspartate receptor co-agonist, in adult brain. *J Biol Chem, 285*(53), 41380-41390. doi:10.1074/jbc.M110.187443

Yang, L., Yang, J. L., Byrne, S., Pan, J., & Church, G. M. (2014). CRISPR/Cas9-Directed Genome Editing of Cultured Cells. *Curr Protoc Mol Biol, 107*, 31.31.31-17. doi:10.1002/0471142727.mb3101s107

Yoshida, K., Furuya, S., Osuka, S., Mitoma, J., Shinoda, Y., Watanabe, M., . . . Hirabayashi, Y. (2004). Targeted disruption of the mouse 3-phosphoglycerate dehydrogenase gene causes severe neurodevelopmental defects and results in embryonic lethality. *J Biol Chem, 279*(5), 3573-3577. doi:10.1074/jbc.C300507200

Zambrowicz, B. P., Imamoto, A., Fiering, S., Herzenberg, L. A., Kerr, W. G., & Soriano, P. (1997). Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. *Proc Natl Acad Sci U S A, 94*(8), 3789-3794. doi:10.1073/pnas.94.8.3789

Zheng, N., Li, L., & Wang, X. (2020). Molecular mechanisms, off-target activities, and clinical potentials of genome editing systems. *Clin Transl Med, 10*(1), 412-426. doi:10.1002/ctm2.34

7 Acknowledgement

I would like to take the opportunity to thank my supervisor, Professor Martin Zenker, for his trust and his unlimited support for the success of this research.

I would also like to thank Dr. Sönke Weinert from the Department of Internal Medicine, Division of Cardiology and Angiology, who did not spare with any advice and gave a lot of his time to support the research and complete it in an optimal manner.

I would like to thank all my colleagues in Institute of Human Genetics, in particular Dr. Denny Schanze. I would like to thank also all my colleagues in the Department of Internal Medicine, Division of Cardiology and Angiology (OVGU).

I will not forget to thank Dr. Julia Dahlmann and Dr. George Kensah from the Heart and Thoracisc Surgery, who did not skimp on giving advice to achieve a part of this research.

Many thanks to all the collaborator persons, which they provided me with the patients.

The first and last thanks to my family, who suffered with me through the difficulty of the way and were the psychological support for me in all my steps.

8 Publications

Abdelfattah, F; Kariminejad; A, Kahlert, A.k, Morrison, P, J, Gumus, E, Mathews, K, D, Schanze, D (2020) Expanding the genotypic and phenotypic spectrum of severe serine biosynthesis disorders. *Hum Mutat*, 41(9), 1615-1628-doi: 10.1002/humu 24067.

Abdelfattah, F; *et al.* (under preparation) Functional effects of *PHGDH* mutations associated with serine biosynthesis defects of different severity

9 Erklärung

Hiermit erkläre ich, dass ich die von mir eingereichte Dissertation zu dem Thema “Monogenic causes of severe fetal abnormalities leading to prenatal or perinatal lethality: lessons from Neu-Laxova syndrome” selbständig verfasst, nicht schon als Dissertation verwendet habe und die benutzten Hilfsmittel und Quellen vollständig angegeben wurden.

Weiterhin erkläre ich, dass ich weder diese noch eine andere Arbeit zur Erlangung des akademischen Grades doctor rerum naturalium (Dr. rer. nat.) an anderen Einrichtungen eingereicht habe.

Magdeburg, den 19.12.2022 M.Sc. Fatima Abdelfattah

Appendix A

Appendix A Table 1: Lab equipments.

Product	**Company**
-150°C freezer	PHC Europe B.V.; Etten-Leur, Netherlands
2720 Thermal Cycler, Applied Biosystems	Thermo Fisher Scientific Inc.; Darmstadt, Germany
3500xL Genetic AnalyZer, Applied Biosystems	Thermo Fisher Scientific Inc.; Darmstadt, Germany
4200 TapeStation system	Agilent Technologies Deutschland GmbH; Waldbronn, Germany
8-channel pipet m10, 0.5-10 µl	Sartorius AG; Goettingen, Germany
Agencourt SPRIPlate 96R ring magnetic plate	Beckman Coulter Inc.; Brea, CA, USA
Analytical balance (ACJ 120 -4M)	Kern & Sohn GmbH; Balingen-Frommern, Germany
Autoclave VX-150	Systec GmbH; Linden, Germany
Beckman Coulter Allegra X-15R	Beckman Coulter Inc.; Brea, CA, USA
Bench-top orbital shaker incubator	PHC Europe B.V.; Etten-Leur, Netherlands
Biological safety laminar flow hood (HERAsafe KSII)	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Biomek NXP Laboratory Automation Workstation	Beckman Coulter Inc.; Brea, CA, USA
Butan /Propan gas Bottle (CV470 Plus)	Camping Gaz (Deutschland) GmbH; Hattersheim am Main, Germany
Cell culture incubator MCO-18 AIC (UV)	PHC Europe B.V.; Etten-Leur, Netherlands
Cell freezing container (Mr. Frosty)	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Cell imaging system (EVOS FL Auto)	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Centrifuge (Allegra X-12R)	Beckman Coulter Inc.; Brea, CA, USA
Centrifuge 5415C	Eppendorf SE; Hambug, Germany
CFX96 Real-Time PCR C1000 Touch thermal cycler	Bio-Rad Laboratories GmbH; Feldkirchen, Germany
Chemagic Magnetic Separation Module I	PerkinElmer chemagen Technologie GmbH; Baesweiler, Germany
Chemiluminescence and fluorescence ECL Imager	Intas Science Imaging Instruments GmbH; Göttingen, Germany
Chemiluminescence imaging system (FluorChem 5500)	Alpha Innotech Corporation; San Leandro, USA
CO2 incubator (MCO-19 A/C (UV))	PHC Europe B.V.; Etten-Leur, Netherlands
Cooled benchtop centrifuge (Rotina 420 R)	Andreas Hettich GmbH & Co. KG; Tuttlingen, Germany
Cooled incubator (MIR-154-PE)	PHC Europe B.V.; Etten-Leur, Netherlands

Digital refrigerated microcentrifuge (Perfect Spin)	PEQLAB Biotechnologie GmbH; Erlangen, Germany
Drying and Heating cabinets	BINDER Labortechnik GmbH; Tuttlingen, Germany
Drying oven	Memmert GmbH + Co. KG, Schwabach, Germany
Drying oven (Heraeus)	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Drying ovens with forced convection, VENTI-Line® Prime	VWR International GmbH; Darmstadt, Germany
E143 Electrophoresis Power Supply	Consort BVBA; Turnhout, Belgium
Electrophoresis accessories (gel trays, combs, casting chambers)	PEQLAB Biotechnologie GmbH; Erlangen, Germany
Electroporator (Neon transfection system)	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Erlenmeyer flask 500 ml, SIMAX	Kavalierglass a.s.; Prague, Czech Republic
Flow cytometer (FACS Canto II)	Becton Dickinson GmbH; Heidelberg, Germany
Flow cytometry cell sorting (AriaIII)	Becton Dickinson GmbH; Heidelberg, Germany
Fluorescence microscope	Carl Zeiss Microscopy GmbH; Jena, Germany
Fluorescence microscope Axiocam ERc5s	Carl Zeiss Microscopy GmbH; Jena, Germany
Freezer and fridges (4°C, -20°C and -80°C)	Liebherr-International S.A.; Bulle, Switzerland
	Robert Bosch GmbH; Gerlingen, Germany
	PHC Europe B.V.; Etten-Leur, Netherlands
Gel-documentation system	PEQLAB Biotechnologie GmbH; Erlangen, Germany
Hemocytometer (Neubauer)	Paul Marienfeld GmbH & Co. KG; Lauda-Königshofen, Germany
Heraeus Labofuge 400	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Heraeus Pico 17 Centrifuge	Thermo Fisher Scientific Inc.; Darmstadt, Germany
HERAsafe KSP 18 Microbiological Safety Cabinet	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Ice machine ZBE 30-10	Intercontinentale ZIEGRA Eismaschinen GmbH; Isernhagen, Germany
iCycler	Bio-Rad Laboratories Inc.; Hercules, CA, USA
IKA magnetic stirrer with heating, RH digital	IKA®-Werke GmbH & CO. KG; Staufen, Germany
IKA® MS 1 mini shakers	IKA®-Werke GmbH & CO. KG; Staufen, Germany

IKA® MS 3 basic shakers	IKA®-Werke GmbH & CO. KG; Staufen, Germany
Incubation hood TH15	Edmund Bühler GmbH; Bodelshausen, Germany
inoLab pH meter	Xylem Analytics Germany Sales GmbH & Co. KG; Weilheim, Germany
Vacusafe™ Comfort Aspiration system	INTEGRA Biosciences GmbH; Biebertal, Germany
CORIO CD-BC4 circulation thermostat	JULABO GmbH; Seelbach, Germany
SW22 Shaking waterbath	JULABO GmbH; Seelbach, Germany
Laboklav 80V autoclave	SHP Steriltechnik AG; Detzel Schloss, Germany
Laboratory microwave	AEG Hausgeräte GmbH; Rothenburg ob der Tauber, Germany
M200 microplate reader	Tecan Group Ltd.; Männerdorf, Switzerland
Masterflex Console Drive	Cole-Parmer Instrument Company Ltd.; Vernon Hills, IL, USA
Masterflex Easy-Load II	Cole-Parmer Instrument Company Ltd.; Vernon Hills, IL, USA
Masterflex PharMed Tubing 24/36	Cole-Parmer Instrument Company Ltd.; Vernon Hills, IL, USA
Measuring cup 2000 ml, plastic	VITLAB GmbH; Grossostheim, Germany
Measuring cylinder 100 ml, plastic	VITLAB GmbH; Grossostheim, Germany
Mettler ToledoAB 104-S/PH Scale	Mettler-Toledo GmbH; Gießen, Germany
Microscope (Evos XL Core)	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Microwave oven R-939	Sharp K.K.; Osaka, Japan
Milli-Q Reference Water Purification System	Merck KGaA; Darmstadt, Germany
Mini centrifuge (Sprout 12V)	Heathrow Scientific; Vernon Hills, IL, USA
Mini-PROTEAN® Tetra Cell	Bio-Rad Laboratories GmbH; Feldkirchen, Germany
Mini-PROTEAN® trans-blot cell	Bio-Rad Laboratories GmbH; Feldkirchen, Germany
MiniSpin centrifuge	Eppendorf SE; Hambug, Germany
Mini-Sub Cell GT Horizontal Electrophoresis System and PowerPac Basic Power Supply	Bio-Rad Laboratories GmbH; Feldkirchen, Germany
Multipette plus/stream	Eppendorf SE; Hambug, Germany
NanoDrop 2000/2000c UV-Vis spectrophotometer	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Ohaus Scout™ Pro Portable Electronic Balance, 400g capacity, 0.1g readability	OHAUS Europe GmbH; Nänikon, Switzerland
PCR thermal cycler (PTC-0200 Gradient Peltier thermal cycler)	MJ Research, Inc.; Waltham, MA, USA

PerfectBlue Gel System Mini S/L	PEQLAB Biotechnologie GmbH; Erlangen, Germany
Phase contrast microscope Axiovert 40 C	Carl Zeiss Microscopy GmbH; Jena, Germany
Pipetting aid (Pipetus®-akku)	Hirschmann Laborgeräte GmbH & Co. KG; Eberstadt, Germany
Platform shakers (Titramax101)	Heidolph Instruments GmbH & Co. KG; Schwabach, Germany
Polymax 1040 (5° tilt angle) shaking and mixing devices	Heidolph Instruments GmbH & Co. KG; Schwabach, Germany
Polypropylene Bottle for Vacusafe™ Aspiration System	INTEGRA Biosciences GmbH; Biebertal, Germany
Precision scale PEJ 4200-2M	Kern & Sohn GmbH; Balingen, Germany
QIAcube robotic workstation	QIAGEN N.V.; Venlo, Netherlands
Reagent bottle with screw cap, SIMAX; 250 ml, 500 ml, 1000 ml	Kavalierglass a.s.; Prague, Czech Republic
Refrigerated centrifuge 3K 12	Sigma Laborzentrifugen GmbH; Osterode am Harz, Germany
Refrigerated centrifuge 5417R	Eppendorf SE; Hambug, Germany
Research® plus 10 µL, 20 µL, 100 µL, 200 µL, 1000 µL	Eppendorf SE; Hambug, Germany
Roll mixer RM5	Karl Hecht GmbH & Co KG; Sondheim vor der Rhön, Germany
Stuart SRT6 roller mixer	Cole-Parmer Instrument Company Ltd.; Vernon Hills, IL, USA
Spinbar Magnetic Stirring Bars	Sarstedt AG & Co; Nümbrecht, Germany
Synergy™ HT Multi-detection microplate reader	BioTek Instruments, Inc.; Friedrichshall, Germany
Thermal Printer P93D	Mitsubishi Electric Corporation K.K.; Tokyo, Japan
ThermoCell MixingBlock MB-102	Hangzhou Bioer Technology Co. Ltd.; Hangzhou, China
Thermomixer compact (1.5 ml, 2 ml)	Eppendorf SE; Hambug, Germany
Turbo grill Microwave (Privileg 8020E)	Privileg Bauknecht Hausgeräte GmbH; Stuttgart, Germany
UV-Transilluminator	Decon Science Tec GmbH; Hohengandern, Germany
UV/Vis Spectrophotometer (Nanophotometer P-330)	Implen GmbH; München, Germany
Vacuum pump AQUASTOP	Van der Heijden Labortechnik GmbH; Dörentrup, Germany
Vacuum controller (CVC 3000)	VACUUBRAND GMBH + CO KG; Wertheim, Germany
Vacuum hand operator (VACUBOY)	INTEGRA Biosciences GmbH; Biebertal, Germany

Variable speed Vortex -2 Genie	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Veriti 96-Well Fast Thermal Cycler, Applied Biosystems	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Vortex - Shaker (LAB Dancer)	IKA®-Werke GmbH & CO. KG; Staufen, Germany
Vortex Genie 2	Scientific Industries, Inc.; Bohemia, NY, USA
Vortex Stirrer (REAX-Control)	Heidolph Instruments GmbH & Co. KG; Schwabach, Germany
VWR Duo Cycler	VWR International GmbH; Darmstadt, Germany
VWR® UV-Transilluminators (GenoView)	VWR International GmbH; Darmstadt, Germany
Wash machine	Miele & Cie. KG; Gütersloh, Germany
Water bath (AQUAline AL 25)	Lauda Scientific GmbH; Lauda-Königshofen, Germany
Water purification (Milli-Q Reference)	Merck KGaA; Darmstadt, Germany
Water purification system (TKA)	TKA Wasseraufbereitungssysteme GmbH; Niederelbert, Germany

Appendix B

Appendix B Table 1: Consumables.

Product	**Company**
10 cm culture dishes for LB agar cultures	Greiner Bio-One GmbH; Frickenhausen, Germany
100 mm cell culture dish	Greiner Bio-One GmbH; Frickenhausen, Germany
12 well cell culture plate (Nunclon™ Delta Surface)	Thermo Fisher Scientific Inc.; Darmstadt, Germany
150 mm cell culture dish	Sarstedt AG & Co. KG; Nümbrecht, Germany
24 well cell culture plate (Nunclon™ Delta Surface)	Thermo Fisher Scientific Inc.; Darmstadt, Germany
48 well cell culture plate (Nunclon™ Delta Surface)	Thermo Fisher Scientific Inc.; Darmstadt, Germany
6 well cell culture plate (Nunclon™ Delta Surface)	Thermo Fisher Scientific Inc.; Darmstadt, Germany
96 well cell culture plate (Nunclon™ Delta Surface)	Thermo Fisher Scientific Inc.; Darmstadt, Germany
96 well PCR Plate, non-skirted	Azenta Life Sciences Inc; Waltham, US
Acetate Foil for 96 Well Plate	Sarstedt AG & Co. KG; Nümbrecht, Germany
Alu-Folie	Carl Roth GmbH + Co. KG; Karlsruhe, Germany
Bottle top filter	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Capillary tips XL, 20 µl, crystal	Biozym Scientific GmbH; Hessisch Oldendorf, Germany
Capillary tips, 200 µl	Biozym Scientific GmbH; Hessisch Oldendorf, Germany
Cell scraper	Greiner Bio-One GmbH; Frickenhausen, Germany
Cell strainer	Corning Inc-; Corning, NY, USA
Centrifuge tubes (15 ml and 50 ml)	Greiner Bio-One GmbH; Frickenhausen, Germany
Combitips advanced, 0.1 ml, 0.2 ml, 5 ml	Eppendorf SE; Hambug, Germany
Combitips plus, 0.5 ml, 1.0 ml	Eppendorf SE; Hambug, Germany
Cryovials (Cryo.S) 2 ml	Greiner Bio-One GmbH; Frickenhausen, Germany
Dispense Parafilm	Carl Roth GmbH + Co. KG; Karlsruhe, Germany
Disposable Pasteur pipettes	Carl Roth GmbH + Co. KG; Karlsruhe, Germany
Disposable pastur pipettes	Corning Inc.; Corning, NY, USA

Disposable Scalpel stainless steel	FEATHER Safety Razor Co., Ltd.; Osaka, Japan
Disposal Bags	Carl Roth GmbH + Co. KG; Karlsruhe, Germany
DURAN® laboratory bottle (50 ml, 250 ml, 500 ml,1000 ml)	DWK Life Sciences GmbH; Wertheim/Main, Germany
Erlenmeyer flask (100 ml, 250 ml, 500 ml, 1000 ml)	DWK Life Sciences GmbH; Wertheim/Main, Germany
	Kavalierglass a.s.; Prague, Czech Republic
Filter Tips and Filter Tips wide-bore; 200 µl, 1.000 µl	QIAGEN N.V.; Venlo, Netherlands
Flow cytometer tube	Sarstedt AG & Co. KG; Nümbrecht, Germany
Glass Pasteur pipettes	Hilgenberg GmbH; Malsfeld, Germany
Immobilon®-P Transfer Membrane (PVDF,pore size 0,45 µm)	Merck KGaA; Darmstadt, Germany
Kimtech Science precision wipes	Kimberly-Clark Corp.; Inving, TX, USA
MaiMed solution PF, nitrile gloves	MaiMed GmbH; Neuenkirchen, Germany
MicroAmp 8-Cap Strip, Applied Biosystems	Thermo Fisher Scientific Inc.; Darmstadt, Germany
MicroAmp Fast 96 well Reaction Plate (0.1 mL), Applied Biosystems	Thermo Fisher Scientific Inc.; Darmstadt, Germany
MicroAmp Fast Reaction Tubes (8 Tubes/Strip), Applied Biosystems	Thermo Fisher Scientific Inc.; Darmstadt, Germany
MicroAmp Optical 96 Well Reaction Plate, Applied Biosystems	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Microseal® B seal	Bio-Rad Laboratories GmbH; Feldkirchen, Germany
Multiplate® PCR plates™ (Low 96 well Clear)	Bio-Rad Laboratories GmbH; Feldkirchen, Germany
Multiply-µStrip Pro 8-strip	Sarstedt AG & Co. KG; Nümbrecht, Germany
PCR tubes,(RNase-, DNase- and pyrogen-free)	Kisker Biotech GmbH & Co. KG; Steinfurt, Germany
Petri Dish	Greiner Bio-One GmbH; Frickenhausen, Germany
Pipette tips (ep T.I.P.S.®) (20 µl, 200 µl, 1000 µl, 5 ml, 10 ml)	Eppendorf SE; Hamburg, Germany
Plastic bags with inscription pad	neoLab Migge GmbH; Heidelberg, Germany
Reaction tubes with lids (1.5 ml, 2.0 ml)	Eppendorf SE; Hambug, Germany
	Sarstedt AG & Co. KG; Nümbrecht, Germany
Rotilabo®-Blotting papers, thick 0.35 mm	Carl Roth GmbH + Co. KG; Karlsruhe, Germany
safeSeal Tips premium 2.5 µl, steril	Biozym Scientific GmbH; Hessisch Oldendorf, Germany

SafeSeal tubes	Sarstedt AG & Co. KG; Nümbrecht, Germany
Serological pipettes (5 ml, 10 ml, 25 ml, 50 ml)	Sarstedt AG & Co. KG; Nümbrecht, Germany
Standard weighing boats	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Sterile Bulk–25 ml Reservoir Basins (LABCOR)	LTF Labortechnik GmbH & Co. KG; Wasserburg (Bodensee), Germany
Sterile Syringe Filter (0.45 µm, 0.2 µm Cellulose Acetate)	VWR International GmbH; Darmstadt, Germany
Syringe 50 ml	Fresenius Kabi AG; Bad Homburg, Germany
Thermowell Sealing Mats, 96 well	Corning Inc.; Corning, NY, USA
Tips XL, 200 µl, colorless (DNase-, RNase-free)	Biozym Scientific GmbH; Hessisch Oldendorf, Germany

Appendix C

Appendix C Table 1: Chemicals and reagents.

Name of the reagents	**Company**
1 Kb Plus DNA Ladder, invitrogen	Thermo Fisher Scientific Inc.; Darmstadt, Germany
100 bp DNA Ladder, invitrogen	Thermo Fisher Scientific Inc.; Darmstadt, Germany
10x Fast digest green	Thermo Fisher Scientific Inc.; Darmstadt, Germany
10x orange buffer	Thermo Fisher Scientific Inc.; Darmstadt, Germany
10x PWO-DNA polymerase buffer (complete and incomplete)	VWR International GmbH; Darmstadt, Germany
25 mM $MgCl_2$ solution	VWR International GmbH; Darmstadt, Germany
2-Propanol (Isopropanol)	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
6X Orange DNA Loading Dye	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Aceton	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Acrylamide 4K-Solution	AppliChem GmbH; Darmstadt, Germany
Agencourt AMPure XP	Beckman coulter Coulter GmbH; Krefeld, Germany
Agencourt CleanSEQ	Beckman Coulter Inc.; Brea, CA, USA
AKASOLV Aqua Care, with blue colour indicator	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Ammonium persulfate (APS)	Carl Roth GmbH + Co. KG; Karlsruhe, Germany
	SERVA Electrophoresis GmbH; Heidelberg, Germany
Ammonium Sulfate Saturated Solution; BioVision	Abcam PLC; Cambridge, UK
Ampicillin-sodium salt	AppliChem GmbH; Darmstadt, Germany
Aqua, sterile, pyrogen-free water	B. Braun SE, Melsungen, Germany
Betaine solution 5M, PCR reagent; Sigma-Aldrich	Merck KGaA; Darmstadt, Germany
BioScience-Grade, nuclease free and autoclaved water	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Bovine Serum Albumin (BSA)	Biomol GmbH; Hamburg, Germany
Bromophenol blue	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Calciumchlorid Dihydrat	Carl Roth GmbH + Co. KG, Karlsruhe, Germany

Chloroquin-Diphosphat	AppliChem GmbH; Darmstadt, Germany
Ciprofloxacin Hydrochloride; Sigma- Aldrich	Merck KGaA; Darmstadt, Germany
Citric acid	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
cOmplete™ ULTRA Tablets, Mini, EASYpack	Roche Pharma AG; Grenzach-Wyhlen, Germany
Conditioning Reagent, 3500 Series, Applied Biosystems	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Dimethyl Sulfoxide for cell culture (DMSO)	AppliChem GmbH; Darmstadt, Germany
di-Sodium hydrogen phosphate Dihydrate; Sigma-Aldrich	Merck KGaA; Darmstadt, Germany
DMEM (1X) Dulbecco′s Modified Eagle Medium	Thermo Fisher Scientific Inc.; Darmstadt, Germany
DNase I; Sigma-Aldrich	Merck KGaA; Darmstadt, Germany
DPBS (10X) Dulbecco′s Phosphate Buffered Saline	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Ethanol 96%	Otto Fischar GmbH & Co. KG; Saarbrücken, Germany
Ethanol absolute	AppliChem GmbH; Darmstadt, Germany
Ethidium bromide 1%, in H2O	Carl Roth GmbH + Co. KG; Karlsruhe, Germany
	Bio-Rad Laboratories GmbH; Feldkirchen, Germany
Ethylenediaminetetraacetic acid (EDTA) ≥99%	Carl Roth GmbH + Co. KG; Karlsruhe, Germany
	Merck KGaA; Darmstadt, Germany
FastDigest BamHI	Thermo Fisher Scientific Inc.; Darmstadt, Germany
FastDigest EcoRI	Thermo Fisher Scientific Inc.; Darmstadt, Germany
FastDigest HindIII	Thermo Fisher Scientific Inc.; Darmstadt, Germany
FastDigest KpnI	Thermo Fisher Scientific Inc.; Darmstadt, Germany
FastDigest SalI	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Fetal Bovine Serum (FBS)	Thermo Fisher Scientific Inc.; Darmstadt, Germany
	Merck KGaA; Darmstadt, Germany
Genomic DNA Reagents	Agilent Technologies Deutschland GmbH; Waldbronn, Germany
Glycerin	Carl Roth GmbH + Co. KG; Karlsruhe, Germany
Glycine	AppliChem GmbH; Darmstadt, Germany
Glycogen, nuclease-free	Thermo Fisher Scientific Inc.; Darmstadt, Germany

HEPES; Sigma-Aldrich	Merck KGaA; Darmstadt, Germany
HyClone™ Antibiotic/Antimycotic solution (Pen/Strep/Fungiezone) solution	Cytiva Europe GmbH; Freiburg, Germany
Hydrochloric acid (HCl) 5 M	Sigma-Aldrich Corp.; St. Louis, MO, USA
In solution MG-132; Sigma-Aldrich	Merck KGaA; Darmstadt, Germany
Incidin OxyFoam S	Ecolab Inc.; Saint Paul, MN, USA
Kanamycinsulfat	Carl Roth GmbH + Co. KG; Karlsruhe, Germany
LB-Agar (Luria/Miller) for molecular biology	Carl Roth GmbH + Co. KG; Karlsruhe, Germany
LB-Medium (Luria/Miller) for molecular biology	Carl Roth GmbH + Co. KG; Karlsruhe, Germany
L-Glutamine 200 mM(100x); gibco	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Magnesium chloride hexahydrate	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Magnesium sulfate heptahydrate	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Magnesiumchlorid; Sigma-Aldrich	Merck KGaA; Darmstadt, Germany
MEM (1X) Minimum Essential Medium	Thermo Fisher Scientific Inc.; Darmstadt, Germany
MES low moisture content; Sigma-Aldrich	Merck KGaA; Darmstadt, Germany
Methanol Rotipuran® ≥99.0%	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Non-essential Amino Acids MEM NEAA (100X)	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Nonfat dried milk powder	Merck KGaA; Darmstadt, Germany
nuclease free and autoclaved DEPC-Water	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Opti-MEM® I (1X) reduced serum medium	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Paraformaldehyde	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
PBS (1x) Phosphate Buffered Saline	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Phenol/ Chloroform/ isoamyl alcohol	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Platinum™ SuperFi DNA polymerase	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Platinum™ Taq DNA Polymerase	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Platinum™ II Hot-Start Green PCR Master Mix	Thermo Fisher Scientific Inc.; Darmstadt, Germany
POP-7 Performance Optimized Polymer, 3500 Series, Applied Biosystems	Thermo Fisher Scientific Inc.; Darmstadt, Germany

Precision Plus Protein™ Standards (Dual color)	Bio-Rad Laboratories GmbH; Feldkirchen, Germany
Puromycin-Dihydrochlorid	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Pwo DNA Polymerase	VWR International GmbH; Darmstadt, Germany
Resazurin (Sodium salt); Sigma-Aldrich	Merck KGaA; Darmstadt, Germany
RiboLock RNase Inhibitor	Thermo Fisher Scientific Inc.; Darmstadt, Germany
RNaseOUT recombinant ribonuclease inhibitor, invitrogen	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Rotiphorese® 50x TAE Buffer	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Rotiphorese®10x SDS-PAGE	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Rotiphorese®10x TAE-buffer for DNA electrophoresis	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
RQ1DNase stop solution	Promega GmbH; Walldorf, Germany
RQ1RNase-free DNase	Promega GmbH; Walldorf, Germany
SCR7 pyrazine; Sigma-Aldrich	Merck KGaA; Darmstadt, Germany
SeaKem® LE Agarose	Lonza Group AG; Basel, Schweiz
SOB-Medium	Carl Roth GmbH + Co. KG; Karlsruhe, Germany
Sodium chloride	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Sodium dihydrogen-phosphate	SERVA Electrophoresis GmbH; Heidelberg, Germany
Sodium dodecyl sulphate (SDS) 10% solution	Carl Roth GmbH + Co. KG; Karlsruhe, Germany
	AppliChem GmbH; Darmstadt, Germany
Sodium Fluoride minimum 99%, Sigma-Aldrich	Merck KGaA; Darmstadt, Germany
Sodium orthovanadate minimum 90% titration, Sigma-Aldrich	Merck KGaA; Darmstadt, Germany
Sodium pyrophosphate decahydrate, 99%; Sigma-Aldrich	Merck KGaA; Darmstadt, Germany
Sodium-Acetate solution, 3 M, pH 5.2	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Sulfuric acid	Merck KGaA; Darmstadt, Germany
T4 DNA Ligase 100 U (1 U/µl)	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Taq DNA polymerase	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Tetracycline hydrochloride	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Tetramethylethylenediamine (TEMED)	Merck KGaA; Darmstadt, Germany

	SERVA Electrophoresis GmbH; Heidelberg, Germany
Triethylamine	Merck KGaA; Darmstadt, Germany
Tris Hydrochloride for buffer solutions	AppliChem GmbH; Darmstadt, Germany
Tris ultrapure (Tris base)	AppliChem GmbH; Darmstadt, Germany
Tris(hydroxymethyl)aminomethane	Carl Roth GmbH + Co. KG; Karlsruhe, Germany
Triton®X-100; Sigma-Aldrich	Merck KGaA; Darmstadt, Germany
Trypan Blue solution; Sigma-Aldrich	Merck KGaA; Darmstadt, Germany
Trypsin-EDTA; gibco	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Tween®20	Carl Roth GmbH + Co. KG; Karlsruhe, Germany
UltraPure Agarose, invitrogen	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Versene	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Water, nuclease-free	Thermo Fisher Scientific Inc.; Darmstadt, Germany
β-mercaptoethanol ≥99.0%	AppliChem GmbH; Darmstadt, Germany

Appendix C Table 2: The used kits.

Product	**Company**
µMACS™ GFP Isolation Kit	Miltenyi Biotec B.V. & Co. KG; Bergisch Gladbach, Germany
AAVS1 TALE-Nuclease Kit	System Biosciences, LLC; CA, USA
AmpFℓSTR® Identifiler Plus PCR Amplification Kit, Applied Biosystems	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Amplite™ Fluorimetric Total NAD/NADH Assay Kit, AAT Bioquest	Biomol GmbH; Hamburg, Germany
BigDye Terminator v3.1 Cycle Sequencing Kit, Applied Biosystems	Thermo Fisher Scientific Inc.; Darmstadt, Germany
BuccalAmp™ Extraction Kit QuickExtract™ DNA Extraction	Epicentre (an Illumina company) Illumina Inc; California, US
Cell Proliferation ELISA, BrdU (colorimetric)	Roche Pharma AG; Grenzach-Wyhlen, Germany
CellTiter 96® Aqueous Non-Radioactive Cell Proliferation Assay	Promega GmbH; Walldorf, Germany
Chemagic DNA Blood Kit special	PerkinElmer chemagen Technologie GmbH; Baesweiler, Germany
Colorimetric Protein Assay	Bio-Rad Laboratories GmbH; Feldkirchen, Germany
D1000 ScreenTape Assay	Agilent Technologies Deutschland GmbH; Waldbronn, Germany

DC™ Protein Assay Reagent	Bio-Rad Laboratories GmbH; Feldkirchen, Germany
DNA Gel Extraction Kit	Thermo Fisher Scientific Inc.; Darmstadt, Germany
dNTP Mix,10 mM each	Thermo Fisher Scientific Inc.; Darmstadt, Germany
dNTP Set (100 mM), invitrogen	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Expand Long Range dNTPack	Roche Pharma AG; Grenzach-Wyhlen, Germany
FastStart High Fidelity PCR System, dNTPack	Roche Pharma AG; Grenzach-Wyhlen, Germany
GeneJET Gel Extraction Kit	Thermo Fisher Scientific Inc.; Darmstadt, Germany
GeneArt®CRISPR Nuclease OFP Reporter Kit	Thermo Fisher Scientific Inc.; Darmstadt, Germany
GeneArt®Genomic Cleavage Detection Kit	Thermo Fisher Scientific Inc.; Darmstadt, Germany
GeneJET Plasmid Maxiprep Kit	Thermo Fisher Scientific Inc.; Darmstadt, Germany
GeneJET Plasmid Miniprep Kit	Thermo Fisher Scientific Inc.; Darmstadt, Germany
High Capacity cDNA Reverse Transcrition Kit, Applied Biosystems	Thermo Fisher Scientific Inc.; Darmstadt, Germany
illustra GenomiPhi V2 DNA Amplification Kit	GE Healthcare Life Sciences; Little Chalfont, UK
In-Fusion®HD Cloning Kit	Takara Bio Europe SAS; Saint-Germain-en-Laye, France
Invisorb®Spin Blood Mini Kit (250)	Stratec SE; Birkenfeld, Germany
InviTrap®Spin Cell RNA Mini Kit (250)	Stratec SE; Birkenfeld, Germany
Lipofectamine® 3000 Transfection Kit, invitrogen	Thermo Fisher Scientific Inc.; Darmstadt, Germany
MinElute®Gel extraction Kit	QIAGEN N.V.; Venlo, Netherlands
MycoAlert® Mycoplasme Detection Kit	Lonza Group AG; Basel, Switzerland
Neon™ transfection system 100 µl kit, invitrogen	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Phosphoglycerate Dehydrogenase (PHGDH) Activity Assay Kit, BioVision	Abcam PLC; Cambridge, UK
PolyFect® Transfection Reagent	QIAGEN N.V.; Hilden, Germany
QIAamp® DNA Blood Mini Kit	QIAGEN N.V.; Hilden, Germany
QIAamp® DNA FFPE Tissue Kit (50)	QIAGEN N.V.; Hilden, Germany
QIAprep® Spin Miniprep Kit	QIAGEN N.V.; Hilden, Germany
QIAquick® Gel Extraction Kit (250)	QIAGEN N.V.; Hilden, Germany
QuickChange Lightning Multi Site-Directed Mutagenesis Kit	Agilent Technologies Deutschland GmbH; Waldbronn, Germany

RNA ScreenTape Analysis	Agilent Technologies Deutschland GmbH; Waldbronn, Germany
RNaseA, DNase and Protease-free (10 mg /ml)	Thermo Fisher Scientific Inc.; Darmstadt, Germany
RNase-Free Dnase set (50)	QIAGEN N.V.; Hilden, Germany
RNeasy®Mini Kit (50)	QIAGEN N.V.; Hilden, Germany
Super Script ™ III Reverse Transcriptase; invitrogen	Thermo Fisher Scientific Inc.; Darmstadt, Germany
SYBR® Premix Ex Taq™II (Tli RNaseH Plus), Bulk	Takara Bio Europe SAS; Saint-Germain-en-Laye, France
Topo-TA Cloning® Kit	Thermo Fisher Scientific Inc.; Darmstadt, Germany
Western Lightning ® Plus-ECL Enhanced Chemiluminescence Substrate	PerkinElmer chemagen Technologie GmbH; Baesweiler, Germany
ZymoPURE™ II Plasmid Maxiprep Kit	Zymo Research Europe GmbH; Freiburg, Germany

Appendix C Table 3: The prepared buffers.

Name of the prepared buffer	**Components**
Transfection solutions	**2x HBS pH 7.05** 280 mM NaCl 1.5 mM $Na_2HPO_4x2H_2O$ 50 mM Hepes **Calcium chloride ($CaCl_2$)** 2.5 M $CaCl_2x2H_2O$ **Chloroquin** 25 mM
Lysis buffer	1.5% Triton X-100 50 mM HEPES pH 7,6 150 mM NaCl 1.5 mM $MgCl_2$ 10% Glycerin 10 mM sodium pyrophosphate decahydrate ($Na_4P_2O_7x10H_2O$) 5 mM EDTA 4.2 mg/ml sodium fluorid 0.5 mg/ml sodium orthovanadat 1Tab/10 ml Protease- Inhibitor
Collection gel	0.5 M Tris base pH 6.8 (adjust with HCl)
Separating gel	1.5 M Tris base pH 8.8 (adjust with HCl)
Tris-glycine buffer	1.92 M glycine 250 mM Tris base
Blocking buffer	skimmed milk powder, 5% (w/v) Tween-20, 0.1% (v/v) in TBS pH 7.4
Primary Antibody solution	Bovine Serum Albumin BSA, 3% (w/v) Tween-20, 0.1% (v/v) in TBS pH 7.4
Stripping buffers	62.5 mM Tris base pH 6.7 2% SDS add fresh: 100 mM ß-mercaoptoethanol
Washing buffer	**1x TBST** 0.1% Tween20
10x TBS pH 7.6	1.37 M NaCl 160 mM Tris-HCL 40 mM Tris base
SOC medium	It the same Super Optimal Broth medium (SOB) after addition of 20 mM glucose

Appendix D

Appendix D Table 1: Oligonucleotide primers to target exon 7 of PHGDH by constructed GeneArt® CRISPR Nuclease vector.

Primer- name	**Sequence (5'–3')**
gRNA_F	AAACACGTCCAGTGCAGCCCGTTTT
gRNA_R	GGGCTGCACTGGACGTGTTTCGGTG

Appendix D Table 2: Oligonucleotide primers used to amplify housekeeping genes.

Housekeeping gene name	**Primer-name**	**Sequence (5'–3')**
GAPDH	G3PDH-F	TGGTATCGTGGACTCA
	G3PDH-R	ATGCCAGTGACCCGTT
ß-2-microglobulin	ß2-microg-F	CTTTCAGCAAGGACTGGTCTTTC
	ß2-microg-R	TCACATGGTTCACACGGCAG
HPRT1	HPRT1-F	TTGCGACCTTGACCATCTTTG
	HPRT1-R	CTTTGCTGACCTGCTGGATTAC

Appendix D Table 3: Oligonucleotide primers used for cloning and site directed mutagenesis.

Primer- name	**Variant**	**Sequence (5'–3')**
PHGDH_KpnIF	WT	TGCAGTCGACGGTACCGCTTTTGCAAATC
PHGDH_BamHIR		TTCCAAGGATCCCGAGAAGTGGAACTGGAAG
PHGDH 160t_F	c.160C>T	GACTGTGAAGGCCTTATTGTTTGCTCTGCCACCAA
PHGDH 160t_R		TTGGTGGCAGAGCAAACAATAAGGCCTTCACAGTC
PHGDH 488a_F	c.488G>A	GAGAGGTAGCTACCCAGATGCAGTCCTTTGG
PHGDH 488a_R		CCAAAGGACTGCATCTGGGTAGCTACCTCTC
PHGDH 638t_F	c.638C>T	CCTGCCCTCCATGACAGGCTTGC
PHGDH 638t_R		GCAAGCCTGTCATGGAGGGCAGG
PHGDH 781a_F	c.781G>A	CTGCACTGGACATGTTTACGG
PHGDH 781a_R		CCGTAAACATGTCCAGTGCAG
PHGDH 793t_F	c.793G>T	GTTTACGGAATAGCCGCCAC
PHGDH 793t_R		GTGGCGGCTATTCCGTAAAC
PHGDH 793a_F	c.793G>A	GTTTACGGAAAAGCCGCCAC
PHGDH 793a_R		GTGGCGGCTTTTCCGTAAAC
PHGDH 795a_F	c.795G>A	GTTTACGGAAGAACCGCCACG
PHGDH 795a_R		CGTGGCGGTTCTTCCGTAAAC
PHGDH 1129a_F	c.1129G>A	CCCGCAGTCATTGTCAGCCTCCTGAAAGAGG
PHGDH 1129a_R		CCTCTTTCAGGAGGCTGACAATGACTGCGGG
PHGDH 1273a_F	c.1273G>A	GCCTCCTGGCCATGGCCCTGGCA
PHGDH 1273a_R		TGCCAGGGCCATGGCCAGGAGGC

PHGDH 1297t_F	c.1297C>T	GGCGCCCCTTACTAGGCTGTGGGCT
PHGDH 1297t_R		AGCCCACAGCCTAGTAAGGGGCGCC
PHGDH 1468a_F	c.1468G>A	GCAGAGGCAGGCATGCGGCTGCTGT
PHGDH 1468a_R		ACAGCAGCCGCATGCCTGCCTCTGC

Appendix D Table 4: Oligonucleotide primers used for Gibson cloning.

Primer- name	**Sequence (5'–3')**
pAAVS1_PHGDH_In Fusion-Sall F	GCGTGATATCGTCGACTCCAGCAATGGCTTTTGCA
pAAVS1_PHGDH_In Fusion-Sall R	ACAGTCTAGAGTCGACTTAGAAGTGGAACTGGAAGGCT

Appendix D Table 5: Oligonucleotide primers used to sequence cDNA of PHGDH on NM-006623.4.

Primer- name	**Sequence (5'–3')**
PHGDH_cDNAseqF1	GTCATCAACGCAGCTGAGAA
PHGDH_cDNAseqF2	CTCCACGACAGGCTTGCT
PHGDH_cDNAseqF3	AGAAGCTCTGGGGACACTGA
PHGDH_cDNAseqF4	AGACCTCTGACCCTGCAATG
PHGDH_cDNAseqR1	CCGCTATCAGCTCCTCTTTG
PHGDH_cDNAseqR2	AGAATTCCCAGGGTCTTTCC
PHGDH_cDNAseqR3	AGCAATTTCCTCCCCACAG
PHGDH_cDNAseqR4	AAGCCCACAGCCTGGTAAG
PHGDH_cDNAseqR5	ATTGGGGGTGTTCATAACCAAGA

Appendix D Table 6: Oligonucleotide primers used for Junction PCR.

Primer- name	**Sequence (5'–3')**
AAVS1 PC-F left of insert	CCGGAACTCTGCCCTCTAAC
AAVS1 PC-R left of insert	CCCGTGAGTCAAACCGCTAT
AAVS1 PC-F right of insert	AGCTATCTGGTCTCCCTTCC
AAVS1 PC-R right of insert	TCCTGGGATACCCCGAAGAG
AAVS1 MCS-R left of insert (PGK promoter)	CATGCTCCAGACTGCCTTGG

Appendix D Table 7: Oligonucleotide primers used for qPCR.

Primer- name	**Sequence (5'–3')**
PHGDH_cDNA_EX11F	CTGCAATGCTGCCTACCATG
bGH_ pAAVS1R	GATGGCTGGCAACTAGAAGG
PHGDH_cDNA_3`UTR R	GGGACCAGTGAGCTCCAAGG

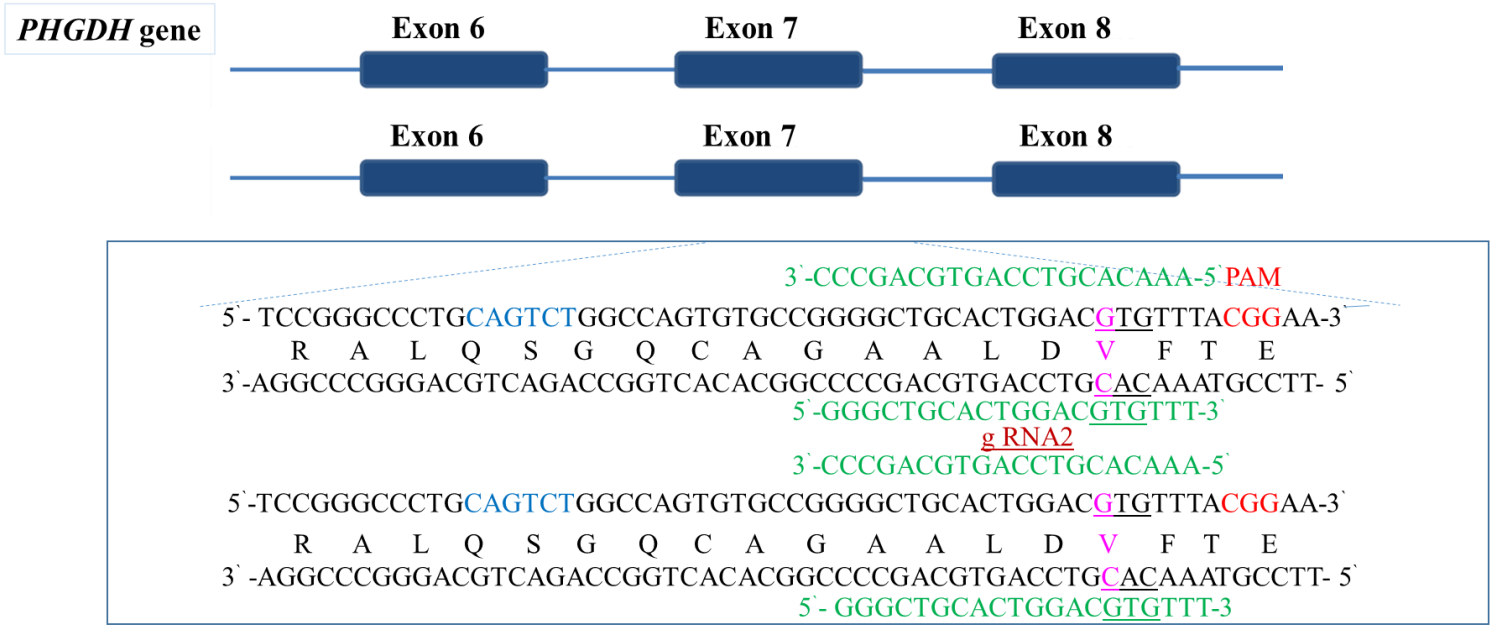
Appendix D Table 8: Oligonucleotide primers for PCR and sequencing.

Primer- name (Gene name_Exon number)	**Sequence (5'–3')**	**Amplicon**
PHGDH_E01F	CGCGGAGAGTTTGAGTATTTC	438 bp
PHGDH_E01R	GCCCCGGAACTAATTGATAC	
PHGDH_E02F	TGAGTCCTAGCCCACATTTC	467 bp
PHGDH_E02R	CTAGACCAGGCATCCTCCTG	
PHGDH_E02F_1	CTGCTTCAGGACTGTGAAGG	130 bp
PHGDH_E02R_1	CCTCCAGATCCACATTGTCC	
PHGDH_E02F_2	GATGTGCCTGCGGCTTTA	110 bp
PHGDH_E02R_2	GGTCACCTTGGTGGCAGA	
PHGDH_E03F	GAGAGAACACAGCCTGCACTC	462 bp
PHGDH_E03R	CAGGGTATGGAGAACATTTGC	
PHGDH_E04+05F	TGTCATGGGCAGTGACTGTG	468 bp
PHGDH_E04+05R	GGCTCCAAGAGAGGGTCAG	
PHGDH_E05F_1	TCCTTTTGCCTGTTTGGTTG	146 bp
PHGDH_E05R_1	CACAGGGTTCCAGCAACATC	
PHGDH_E05F_2	AGCGGAAGAAGGTGAGCAG	144 bp
PHGDH_E05R_2	CCCAGGGTCTTTCCATTCA	
PHGDH_E05F_3	GAAAGACCCTGGGAATTCTTG	135 bp
PHGDH_E05R_3	GCCTTTCCCAAAATTGCTG	
PHGDH_E06F	GCTGAGCATGGTAGTTAGTATATGG	457 bp
PHGDH_E06R	CAGCTTACCTAGACCCCAAGC	
PHGDH_E06F_1	AGAGGTCTCGGCCTCCTTT	152 bp
PHGDH_E06R_1	CTGGTCTGTGGCTTCTGTGA	
PHGDH_E06F_2	TCTGTGATTTCATCACTGTGC	141 bp
PHGDH_E06R_2	TTACCACTGTCTGGGGAAGG	
PHGDH_E07F	GGAGAGAGGCTCTGGGAAAG	469 bp
PHGDH_E07R	CACATAGCTGCCTCCATACC	
PHGDH_E08F	TGAGATGGAACATTCATGTGG	444 bp
PHGDH_E08R	AGCTCCCAGGGCTGAATC	
PHGDH_E08F_1	GTCCATGGCAGCCAACTTAG	148 bp
PHGDH_E08R_1	AGCAATTTCCTCCCCACAG	
PHGDH_E09F	CCTGTTGTTAGTGGCTGATGG	428 bp
PHGDH_E09R	TACTCAGGGCAGGGTCCAAG	
PHGDH_E10F	AGCGAGGAGCCCATAGTCC	437 bp
PHGDH_E10R	CACAAAGCAGTGCCATTAGC	
PHGDH_E11F	CTGGAGCCCTACATCCTACC	446 bp
PHGDH_E11R	ACCCTGGGACATGAAGAGG	
PHGDH_E12F	ACCATCATCCGCTCATTGC	467 bp
PHGDH_E12R	CAAAAGTGCAGAAGGAGTTGG	
PSAT1_E01F	AGGAGCAACTGCTTCGACTC	486 bp
PSAT1_E01R	TCCTCACCAACTGCATTCTTC	
PSAT1_E01F_1	ACTCGGCGCAGGAACAAG	164 bp

PSAT1_E01R_1	GACCAGGCCCAAAGTTGAC	
PSAT1_E01F_2	GTCCTCCTTGGCTGACTCAC	140 bp
PSAT1_E01R_2	CCGCCTGAACCTCACTCC	
PSAT1_E02F	GTTGTCCCTTTAACCCAGAAC	452 bp
PSAT1_E02R	CCTTTGAAACATGGACATTCAC	
PSAT1_E03F	TTCCCTGTATTGTTGTTTACTTGC	495 bp
PSAT1_E03R	TGCTGACAAAGATGACCATTTC	
PSAT1_E03F_1	TCATTTGGTTTTGAGCTGGA	162 bp
PSAT1_E03R_1	AACAAAAGTAACATGATTTGGTCTT	
PSAT1_E03F_2	AGGTCATCAGATTTTGCCAAG	155 bp
PSAT1_E03R_2	TCTGAACAAACAATAGAGTGTAGAAAA	
PSAT1_E04F	AAATCCACCCAAGTTAAAGCAC	444 bp
PSAT1_E04R	AGCTAGTACAGCCACCCAATC	
PSAT1_E04F_1	CCTTAAACCTCATTGGCTTG	148 bp
PSAT1_E04R_1	CAGAACTTACTTGTATAACTCCCAAGT	
PSAT1_E04F_2	CTGTTCCAGACAACTATAAGGTGA	130 bp
PSAT1_E04R_2	GACCAAGCTCCTGTCACCAC	
PSAT1_E04F_3	AAGGTGATTTTTCTGCAAGGA	111 bp
PSAT1_E04R_3	CCAAGCTCCTGTCACCACAT	
PSAT1_E05F	GTTGGATCGCTCATCCTCTC	485 bp
PSAT1_E05R	TTCTGCAGGCCTCAAACAC	
PSAT1_E05F_1	ATCCAAGCACCTGGAACC	152 bp
PSAT1_E05R_1	CACTGGCTTGGACAGGAAGT	
PSAT1_E05F_2	GGGAGCAGTACTGGTTTGTG	107 bp
PSAT1_E05R_2	ACCGAGTCAGTGGGTTCC	
PSAT1_E05F_3	GGAGCAGTACTGGTTTGTGACAT	99 bp
PSAT1_E05R_3	CAGTGGGTTCCCCTCACC	
PSAT1_E05F_4	TTCTGTGATAGAAATTCCAGATCC	105 bp
PSAT1_E05R_4	AAGTCAAACTCCACACCATGC	
PSAT1_E06F	TGAACTGCTCCCTACACAGG	464 bp
PSAT1_E06R	TGACCTCAAGTGGTCCATCC	
PSAT1_E07F	GCTTTCCAACTCCTGGCTTC	474 bp
PSAT1_E07R	CAGGACCACACACACTCACC	
PSAT1_E07F_1	GCAAAGATGAGCTAAACGGATT	135 bp
PSAT1_E07R_1	TTGTTTGAGATTTGATGGAGCTA	
PSAT1_E07F_2	TGAGCCACTGCATTTGACTA	178 bp
PSAT1_E07R_2	AATCCACTCCAGAACCAAGC	
PSAT1_E08F	TTCTGCTTGCATCTAGGAAGTG	436 bp
PSAT1_E08R	TCTCAGCTGGGGAATTATGC	
PSAT1_E08F_1	CATGTTTTTCAGTTGTCCAGTG	170 bp
PSAT1_E08R_1	TGCATTGCAGATGTACTCACC	
PSAT1_E08F_2	AATGCCAAAGGAGATGATGC	155 bp
PSAT1_E08R_2	GATTTGTTTTATTCTGCATTTGAA	
PSAT1_E08F_3	TGCATCTAGGAAGTGTTAAGAAAA	230 bp
PSAT1_E08R_3	CCTTTCAAGGACAACATATTGAG	

PSAT1_E08F_4	TTCTGCTTGCATCTAGGAAGTG	180 bp
PSAT1_E08R_4	GCATCATCTCCTTTGGCATT	
PSAT1_E09F	TCAGGTGCTGCAACTCTCAG	511 bp
PSAT1_E09R	TCCACAGGCAAGTTTGAAATC	
PSPH_E04F	AGATTGGCTTAGTAGGCAAGG	465 bp
PSPH_E04R	GTTAGTGACGGTTGTCATTCC	
PSPH_E05F	TGATCTGCCCACCTCAGC	511 bp
PSPH_E05R	GAAAATAAATAAACCACCCAGAGG	
PSPH_E05F_1	AGGGAGCAGGTGCAGAGA	141 bp
PSPH_E05R_1	AAAGGAATAGTTAGATGCTCTTGG	
PSPH_E05F_2	ATCCAGCCCTCCAGGGAGCAGG	108 bp
PSPH_E05R_2	GGTGCTGAAATACACCTGGAG	
PSPH_E06F	GGCAGTATACCTTGTCAAACATACC	450 bp
PSPH_E06R	TGAGCAAAGGGTACCAAAGG	
PSPH_E06F_1	TTGTATTACTCAGTGTTGGGTATTTT	459 bp
PSPH_E06R_1	AGGGGGCTTCGTAGGAGTTA	
PSPH_E07F	ATTCCCAGGAACAGGTGTGG	467 bp
PSPH_E07R	CCAGGAGGCGGATATTGC	
PSPH_E08F	ACCTGGAGTCACAGCTACTCA	385 bp
PSPH_E08R	AGCAAGTTGTAATAGGCAACTGT	

Appendix E



Allele 1: CAGTCTGGCCAGTGTGCCGGGG--------ACTGGACGTGTTTACGGAA **(c.770_773delCTGC) 4 bp**

Allele 2: CAGTCTGGCCA----------CTGCACTGGACGTGTTTACGGAA **(c.759_769 delGTGTGCCGGGG) 11 bp**

Appendix E Figure 1: Schematic representation of my strategy to target exon 7 of PHGDH gene in wtHEK cells, followed by the sequence of the resulting alleles of the disrupted exon.

Two single stranded DNA oligonucleotides were designed, one encoding crRNA sequence to target a specific site of exon 7 of PHGDH and the other on the reverse strand (It has been called gRNA2). This target sequence is 20 bp in length and it is adjacent to the PAM sequence (CGG), which is on the 3' prime end of the target sequence. After the cloning of the double-stranded DNA oligo encoding the corrresponding site of exon 7 in the gRNA cloning cassette of the vector CRISPR Nuclease OFP reporter, a single guide RNA (gRNA2) was generated to direct the expressed Cas9 nuclease from the same vector to the corresponding genomic location. Cas9-induced DSB are repaired via the NHEJ DNA repair pathway. The repair is error-prone, and thus deletions were introduced that can disrupt gene function. Two deletion alleles as shown in the figure were recognized in the generated koHEK cell line.

Supporting Results related to Appendix E Figures 2, 3, and 4

Structural analysis and molecular modeling

To provide a better understanding of the functional effects of observed missense variants in *PHGDH* and *PSAT1*, structural analysis was performed based on the crystal structures of PHGDH (PDB code: 2G76) and PSAT1 (PDB code: 3E77). Wildtype and mutant residues were modeled with SwissModel (Guex & Peitsch, 1997) and RasMol (Sayle & Milner-White, 1995) was used for structure analysis and visualization.

Structural information is available for the N-terminal parts of the two proteins. PHGDH: amino acids 6-307 (of 533); PSAT1: amino acids 17-370 (of 496). No structure is available for the C-terminal regulatory domain of eukaryotic PHGDHs.

Results

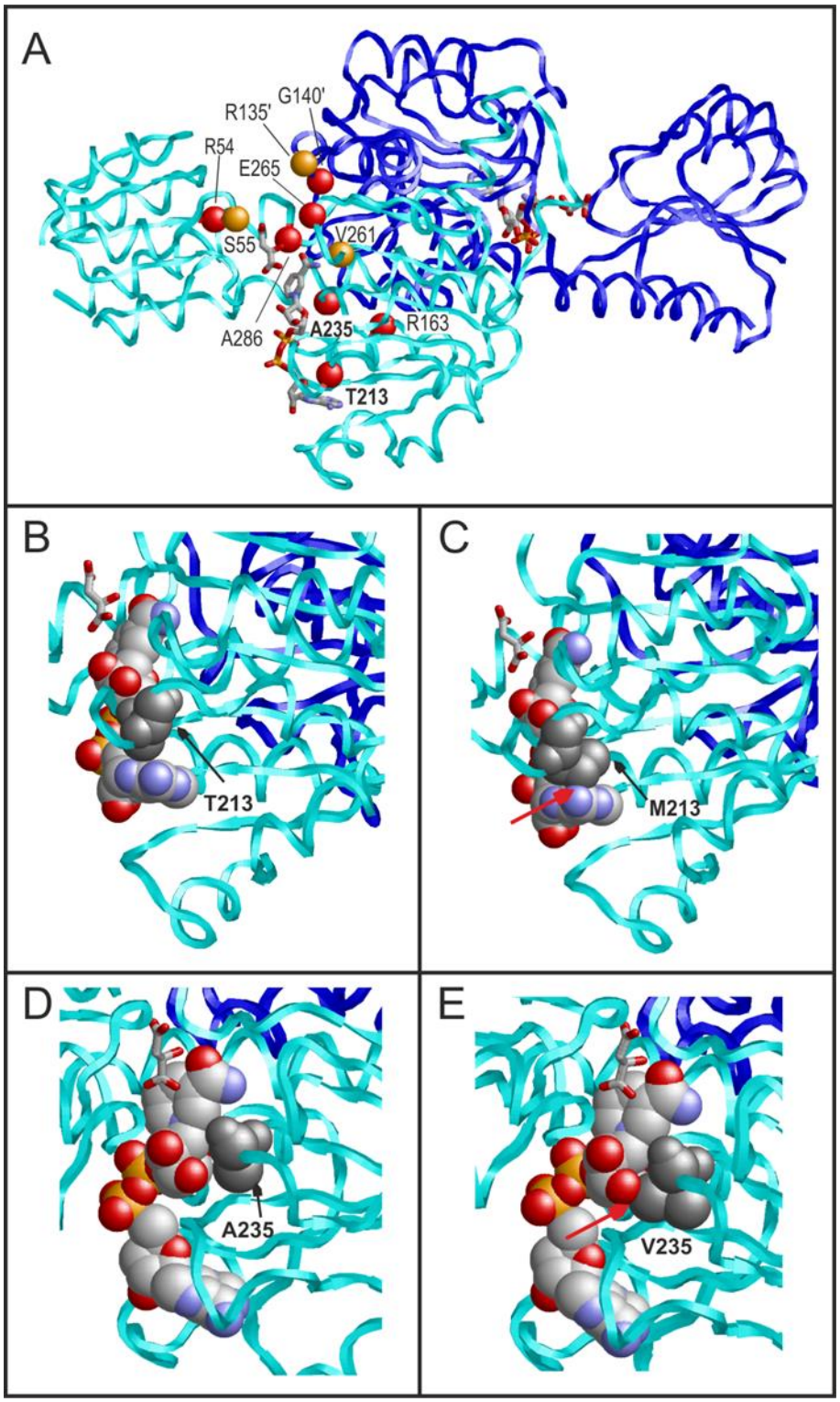
For PHGDH, the novel variants observed in the present study, p.Thr213Met (T213M) and p.Ala235Val (A235V), were particularly analyzed. Both variants are in the immediate vicinity of the NAD^+ cofactor (**Appendix E Figure 2A**). A closer inspection of the interactions revealed that both T213 (**Appendix E Figure 2B**) and A235 (**Appendix E Figure 2D**) form direct interactions with NAD^+ thereby stabilizing the cofactor binding. At both sites, the replacement by bulkier residues results in steric problems ('clashes') that are expected to hamper NAD^+ binding (**Appendix E Figure 2 (C and E)**) thereby probably causing a reduced enzymatic activity. An inspection of the variants previously reported (**Appendix F Table 2**, **Appendix E Figure 2**) reveals that a significant portion of them also cluster around the substrate and cofactor binding site (R54C, S55F, R135W, V261M, E265K, A286P) suggesting that they may act by a similar mechanism, i.e. by causing reduced cofactor or substrate affinity.

The remaining variants observed in the globular PHGDH nucleotide-binding domain (G140R, R163W, R163Q) are located at the dimer interface (**Appendix E Figure 2**) suggesting that they may primarily affect dimer stability. Importantly, the dimer interface was shown to play a crucial role in the overall stability of this protein family and an intact dimer is essentially required for optimal enzymatic activity (Mishra *et al.*, 2012). Therefore, amino acid side-chain changes that are destabilizing the dimer configuration are also expected to cause a reduced enzymatic activity.

Compared to PHGDH, the novel missense variants observed in PSAT1, p.Arg61Trp (R61W), p.Gly79Trp (G79W), p.Glu155Gln (E155Q), and p.Cys245Arg (C245R), are rather distributed over the structure (**Appendix E Figure 3A**). Two of the variants (G79W, E155Q) are located near the pyridoxal 5'-phosphate (PLP) cofactor binding site. A closer inspection of the structure reveals that G79 directly interacts with PLP (**Appendix E Figure 3B**) and that the exchange to a bulky tryptophan causes steric clashes with the PLP and with adjacent amino acids (**Appendix E Figure 3C**). From the structural analysis, the E155Q exchange is also expected to affect PLP binding, although by a slightly different mechanism. In the wildtype, E155 forms two hydrogen bonds to the protein backbone thereby stabilizing the orientation of W107 that exhibits tight stacking interactions with the PLP (**Appendix E Figure 3D**). The different physico-chemical properties of the Q155 sidechain result in the loss of one hydrogen bond which is expected to destabilize the W107 position and consequently its interaction with PLP (**Appendix E Figure 3E**). In addition to the two variants analyzed above, the previously reported S179L variant **(Appendix F Table 3**; **Appendix E Figure 3A**) is also located close to the PLP binding site suggesting similar unfavorable effects on cofactor binding. Interestingly, similar structural effects on cofactor binding have also been observed for the T213M and A235V variants in PHGDH (**Appendix E Figure 2**).

C245 is located close to the PSAT1 dimer interface (**Appendix E Figure 4F**) and a replacement by arginine results in steric clashes with V39 and M42 (**Appendix E Figure 4G**), which probably reduce dimer stability. The region of these clashes is close to the site of the S43R variant (**Appendix E Figure 3A**), which has been previously described as disease-related and was also found in this study (**Appendix F Table 3**). Since previous investigations have shown that the dimeric configuration of the enzyme is essential for function (John, 1995; Mishra *et al.*, 2010), variants that destabilize the dimer are expected to cause a reduced enzymatic activity.

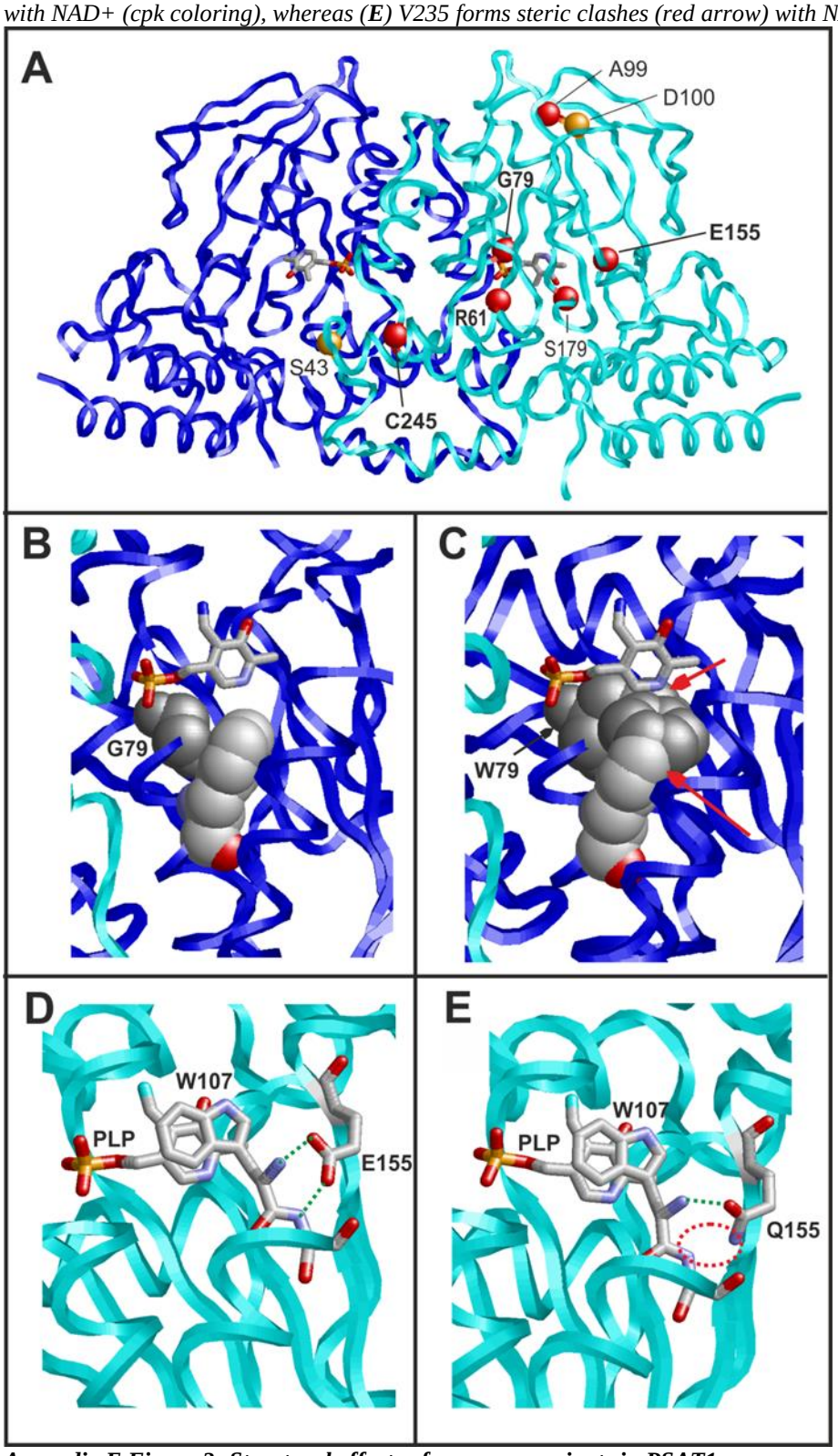
Like the previously described A99V and D100A variants, the newly discovered R61W variant is neither located close to the dimer interface nor to the substrate/cofactor binding site (**Appendix E Figure 3A**) suggesting that these exchanges mainly affect protein/domain stability itself. Molecular modeling indicates that R61W causes a destabilization of the enzyme fold by steric clashes (R61W; **Appendix E Figure 4I**).



Appendix E Figure 2: Structural effects of sequence variants in PHGDH.

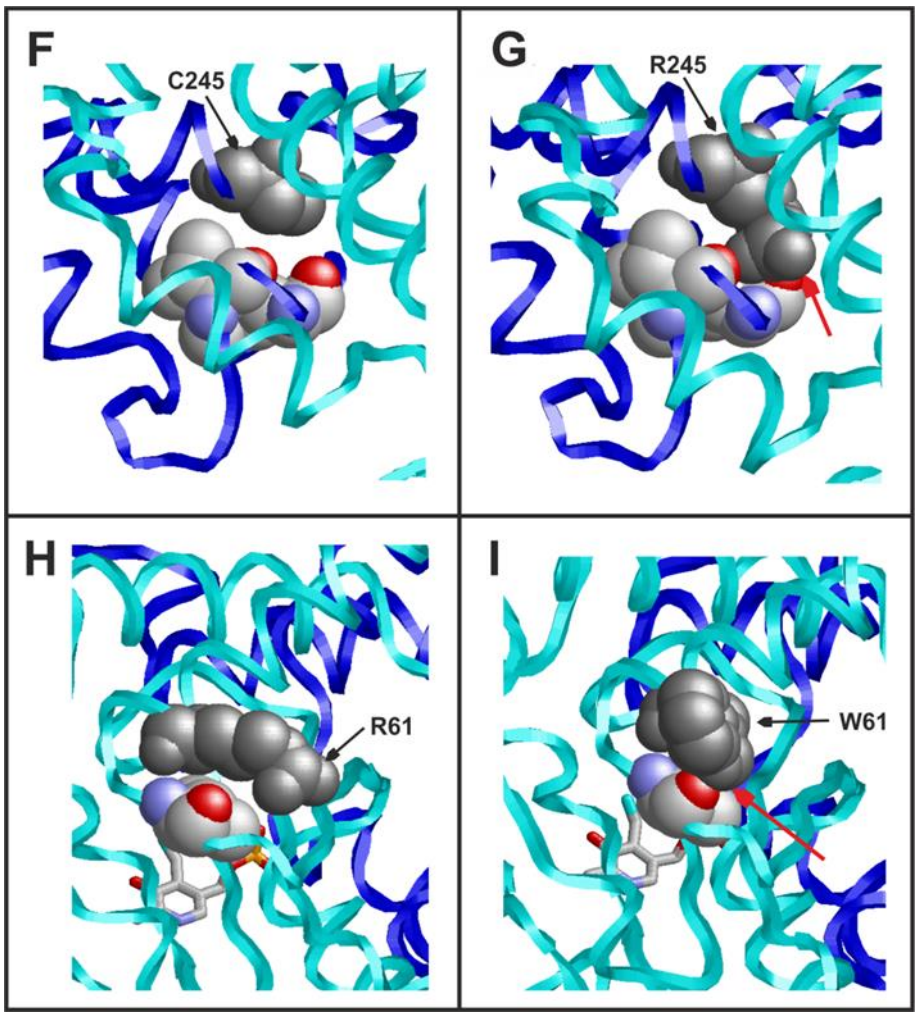
*(**A**) Overview of the dimeric PHGDH structure (cyan, blue) indicating the position of the missense variants. The bound malate and NAD+ are shown in stick presentation and colored according to the atom type (cpk coloring). In one subunit, NLS- and SBDNL-associated variants are highlighted by red and orange balls, respectively. Novel variants detected in the present study are labelled in bold and their structural effect is shown in detail in the following panels. (**B**) T213 (dark grey) forms favorable interactions with NAD+ (space-filled; cpk coloring),*

*whereas (**C**) M213 forms steric clashes (red arrow) with NAD+. (**D**) A235 (dark grey) forms favorable interactions with NAD+ (cpk coloring), whereas (**E**) V235 forms steric clashes (red arrow) with NAD+.*



Appendix E Figure 3: Structural effects of sequence variants in PSAT1.

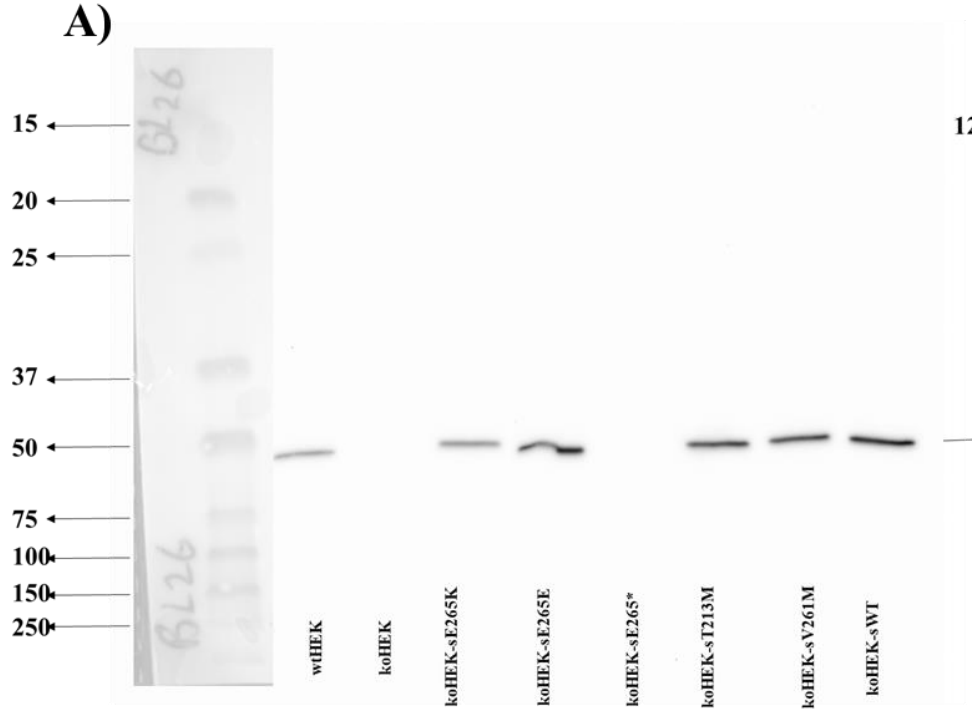
(***A***) *Overview of the dimeric PSAT1 structure (cyan, blue) indicating the position of the missense variants. The bound pyridoxal 5'-phosphate (PLP) is shown in stick presentation and colored according to the atom type (cpk coloring). In one subunit, NLS- and SBDNL-associated variants are highlighted by red and orange balls, respectively. Novel variants detected in the present study are labelled in bold and their structural effect is shown in detail in the following panels. (**B**) G79 (dark grey) forms favorable interactions with PLP (stick presentation; cpk coloring), whereas (**C**) W79 forms steric clashes (red arrows) with PLP and the adjacent F83 (space-filled presentation; cpk coloring). (**D**) E155 forms two backbone hydrogen bonds (green dotted lines) that stabilize the interaction of W107 with PLP. (**E**) Q155 can only form one hydrogen bond resulting in a destabilization of the W107 segment (red dotted circle) and a weakened interaction with PLP. Residues 100-106 have been omitted for clarity in this presentation.*

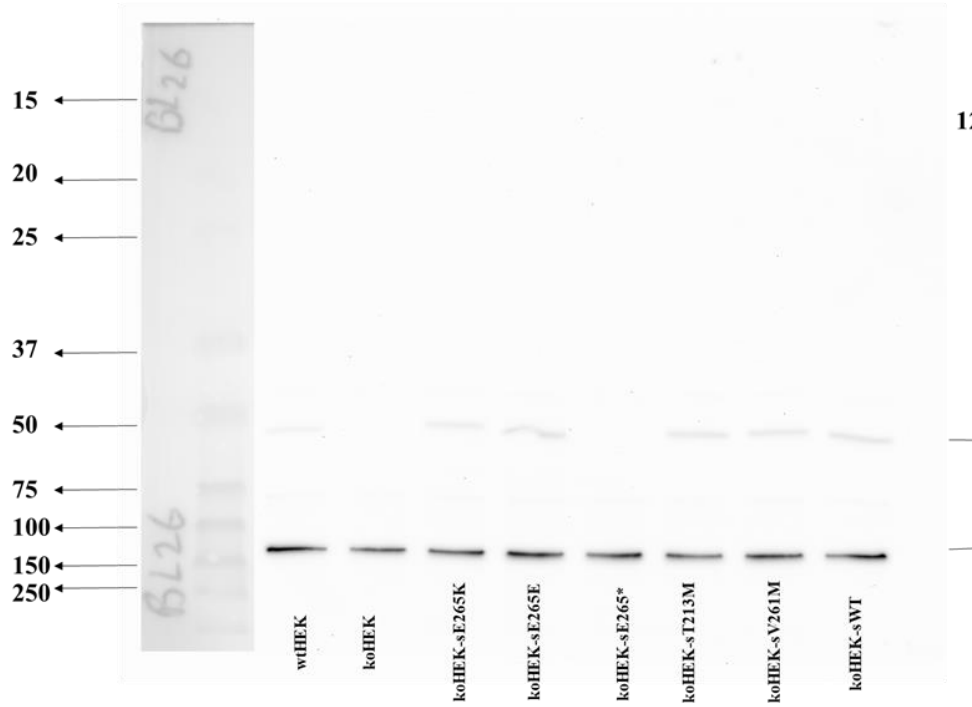


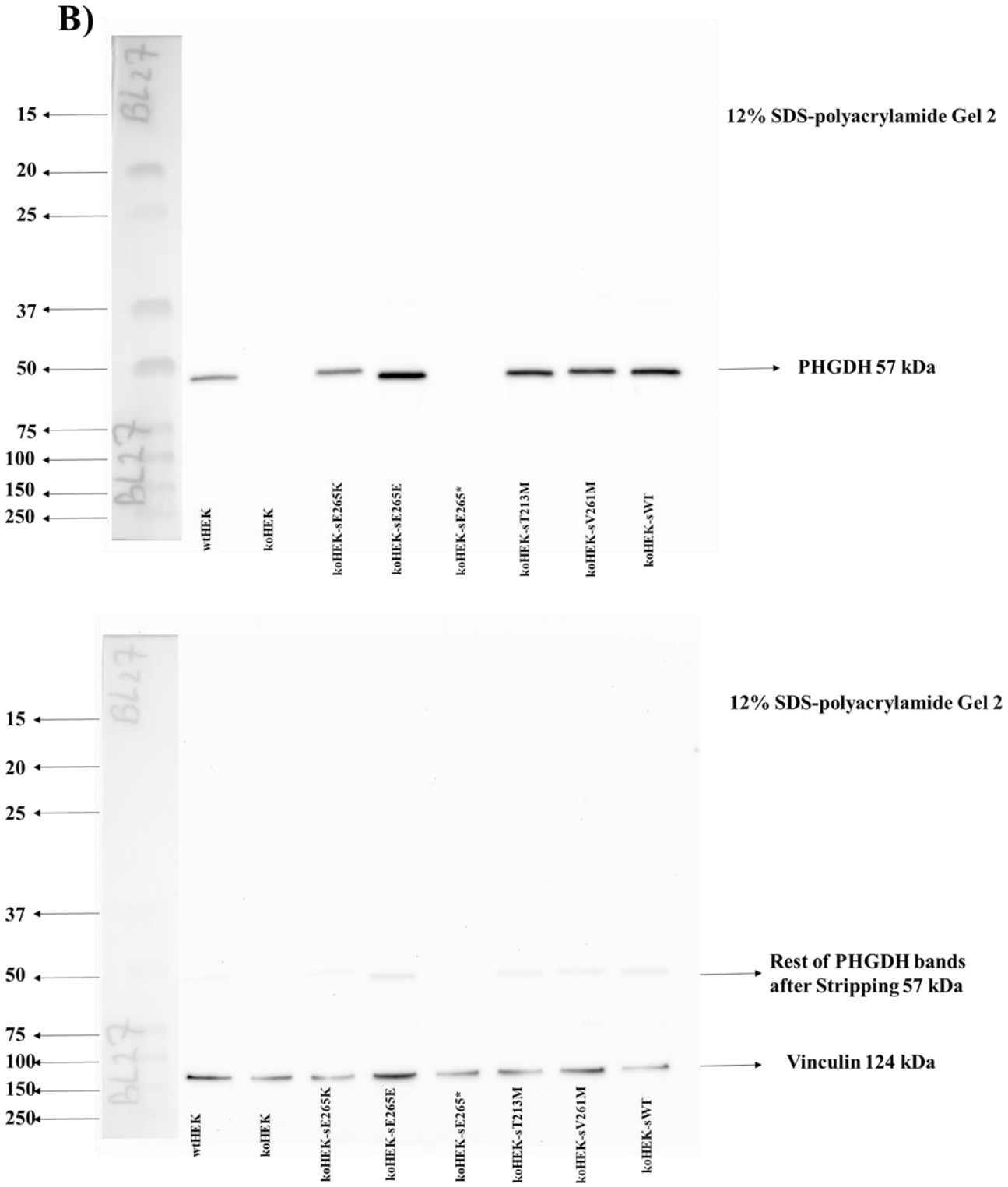
Appendix E Figure 4: Structural effects of sequence variants in PSAT1.

(***F***) *C245 (dark grey) is located close to the dimer interface in the vicinity of V39/M42 (cpk coloring). (**G**) R245 forms clashes with V39/M42 (red arrow) thereby destabilizing the dimer interface. (**H**) R61 (dark grey) is located close to V66 (cpk coloring) but does not form any clashes. (**I**). The bulkier sidechain of W61 causes steric clashes with V66 (red arrow) resulting in a domain destabilization*.

A)

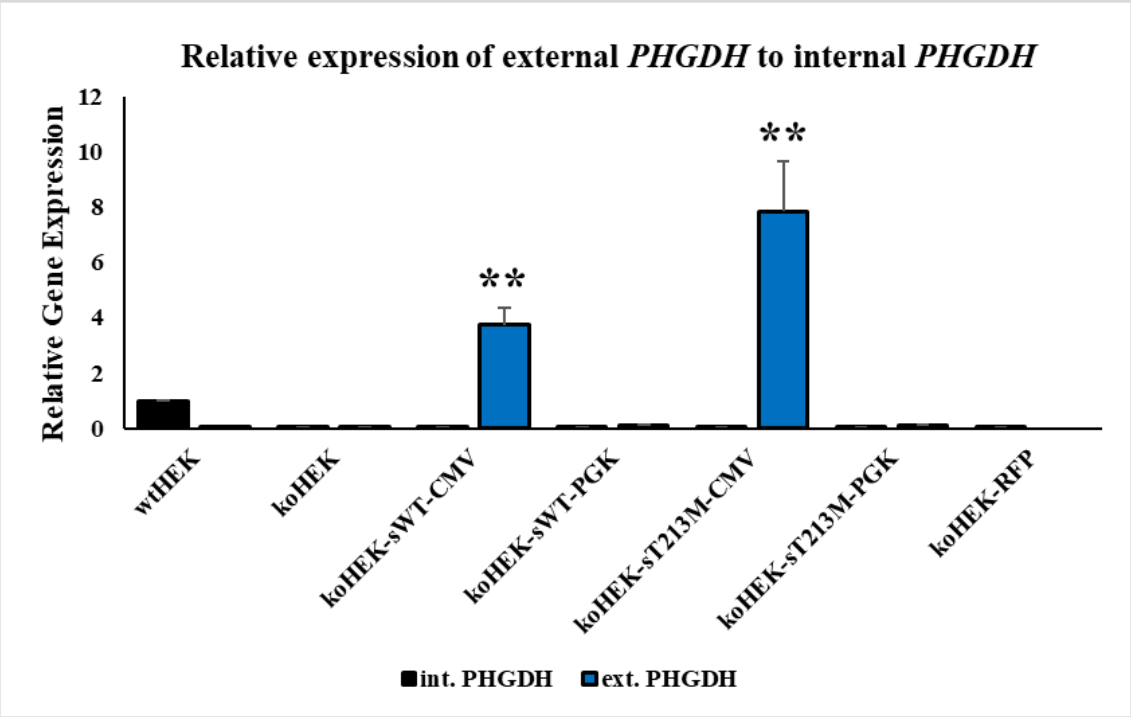

15
20
25
37
50
75
100
150
250
wtHEK
koHEK
koHEK-sE265K
koHEK-sE265E
koHEK-sE265*
koHEK-sT213M
koHEK-sV261M
koHEK-sWT
12% SDS-polyacrylamide Gel 1
PHGDH 57 kDa


15
20
25
37
50
75
100
150
250
wtHEK
koHEK
koHEK-sE265K
koHEK-sE265E
koHEK-sE265*
koHEK-sT213M
koHEK-sV261M
koHEK-sWT
12% SDS-polyacrylamide Gel 1
Rest of PHGDH bands
after stripping 57 kDa
Vinculin 124 kDa



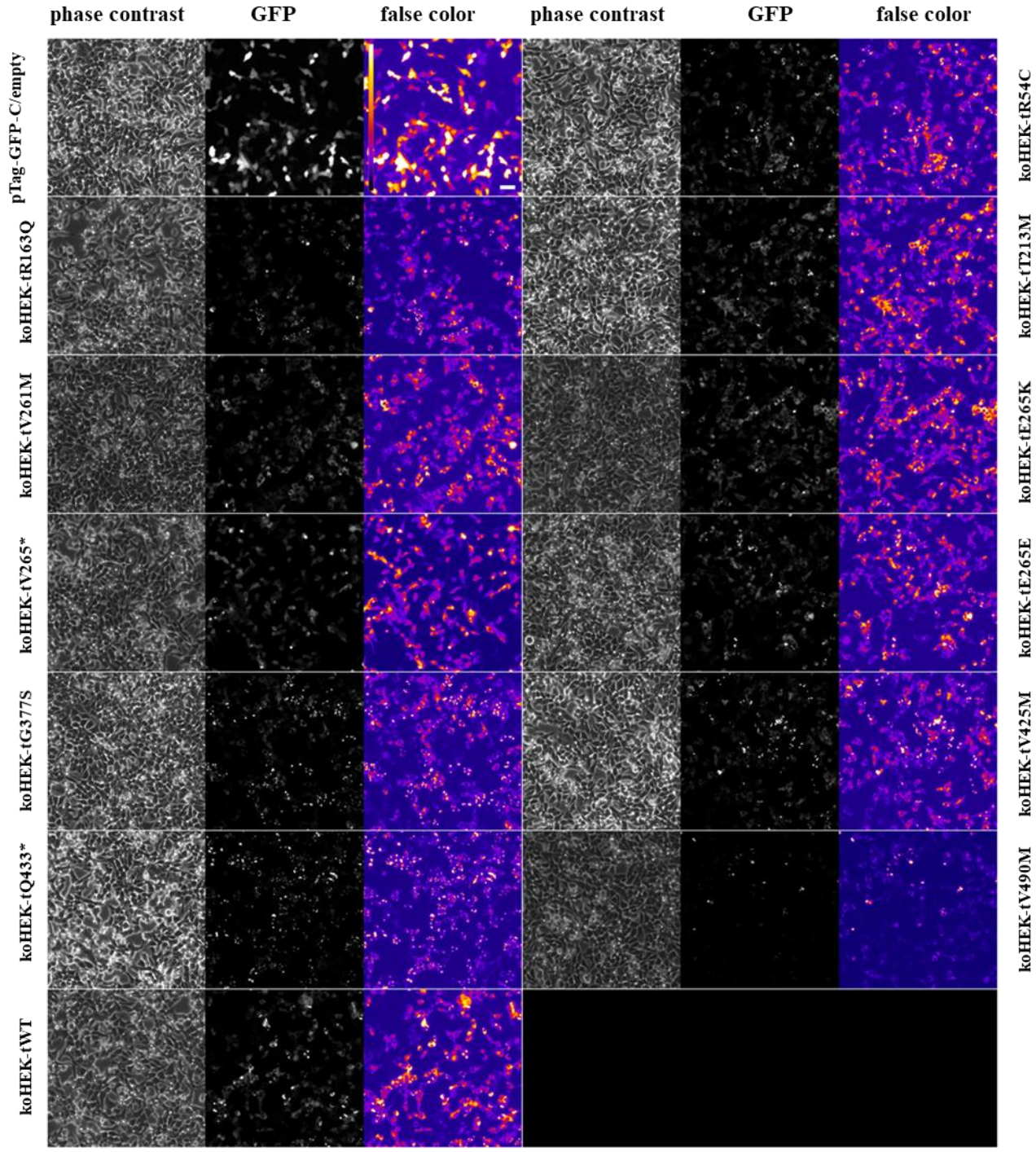
Appendix E Figure 5: (A and B): immunoblotting of PHGDH from knock-in PHGDH HEK cells on SDS-polyacrylamide gels 12%.

The extracted proteins from knock-in HEK cells were loaded on SDS-polyacrylamide gels 12% and the complete gel was shown to check if there is any truncated protein of koHE-sE265: All these genertaed cell lines were with CMV promoter.*



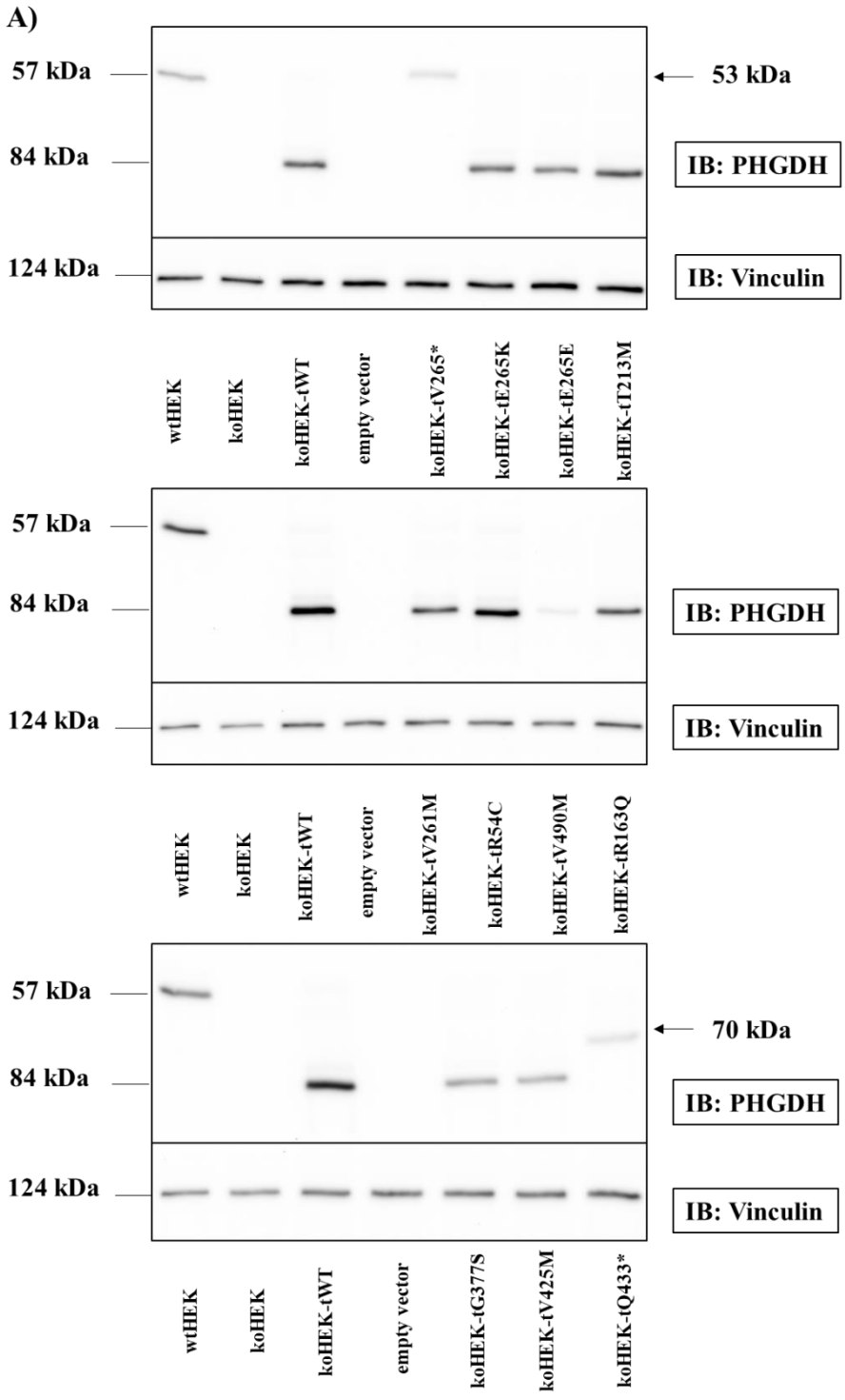
Appendix E Figure 6: Realtive expression of external PHGDH to internal PHGDH in wtHEK, koHEK and knock-in HEK cell lines (qPCR results).

*KoHEK-RFP was included and refers to koHEK co-transfected with Pzt-AAVS1-TALE Nuclease pair and Positive Control Donor vector provided with the kit. Statistical analysis was done by GraphPad InStat using one-way Analysis of Variance (ANOVA) and compared to internal expression of wtHEK cells. Dunnett Multiple Comparisons test **p≤0.01, Data represents Mean of three independent experiments n=3 and the normalization were done to HPRT-1.*



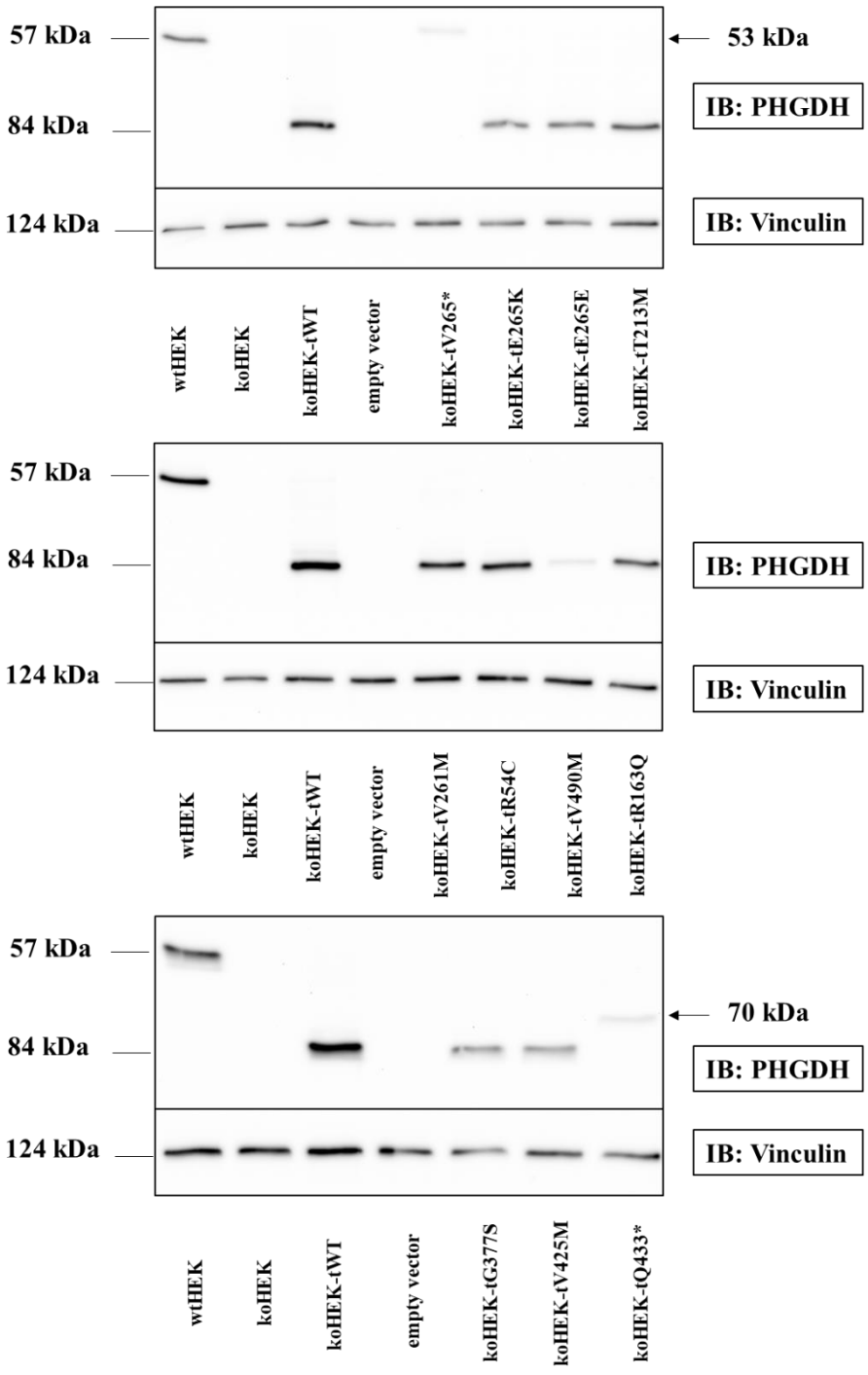
Appendix E Figure 7: Fluorescence microscope images show transfection efficiency in koHEK cells transfected transiently with 11 different variants of PHGDH, beside to WT and empty vector. The false color looks up table was used to visualize intensity differences as colors.

All the images of the transiently transfected koHEK with 11 different variants of PHGDH, beside to wt and empty vector showed almost the same transfection effeciency. The variant which expresses the mutated protein p. Val490Met from koHEK-tV490M and which didnot express the same abundance of the protein variants by immunoblotting of PHGDH showed also almost the same transfection effeciency under the fluorescence microscope.



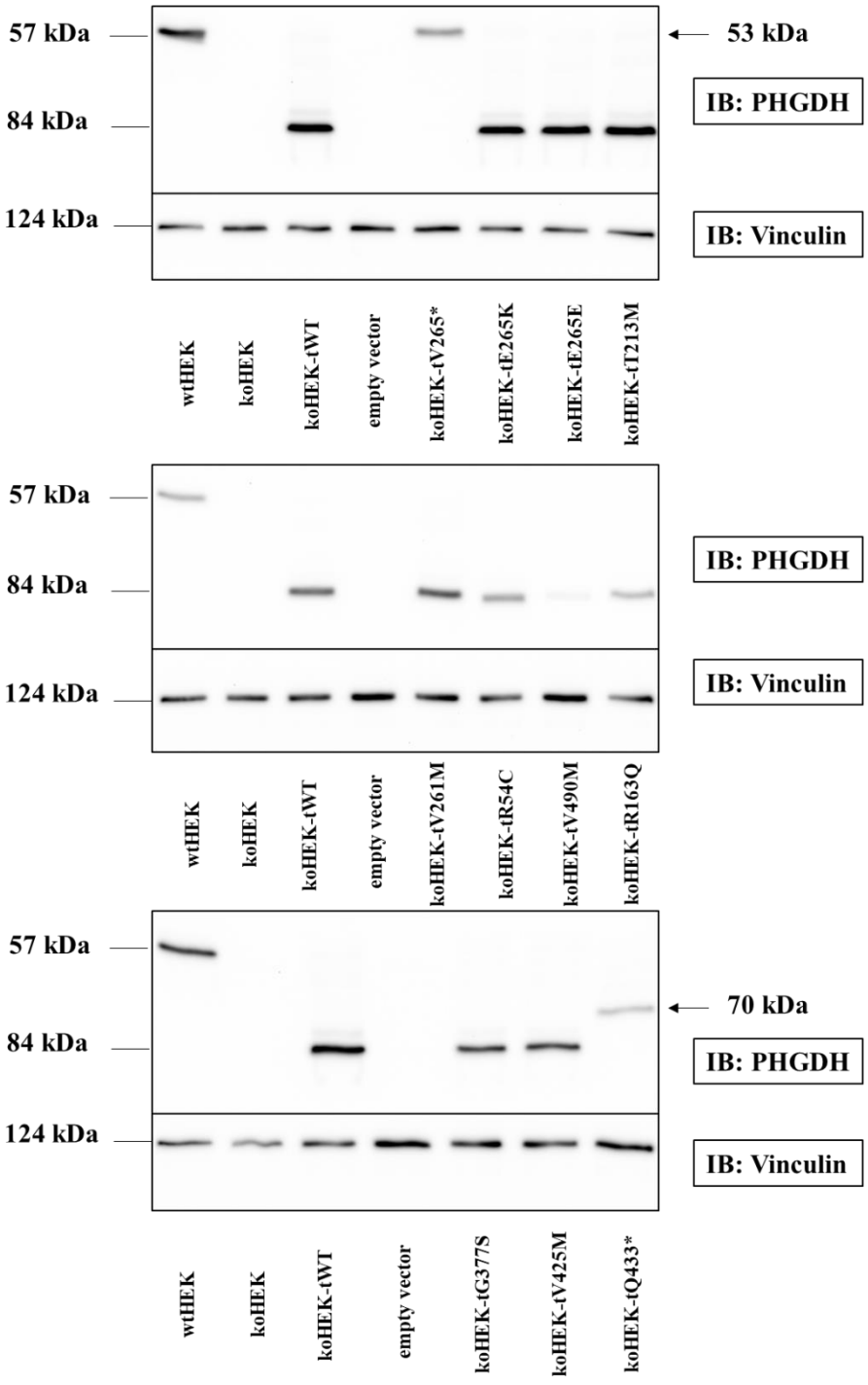
First experiment

B)

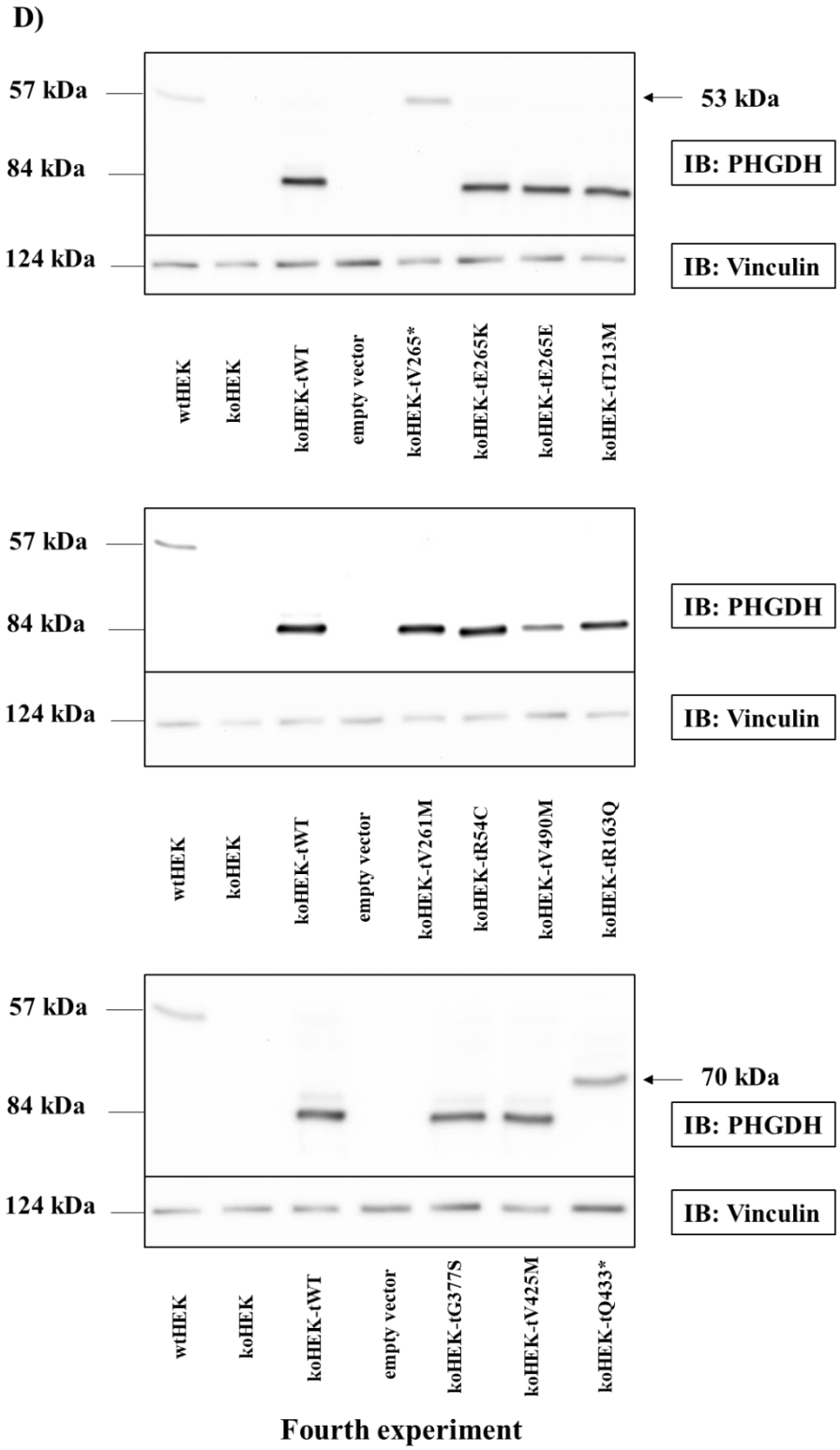


Second experiment

C)



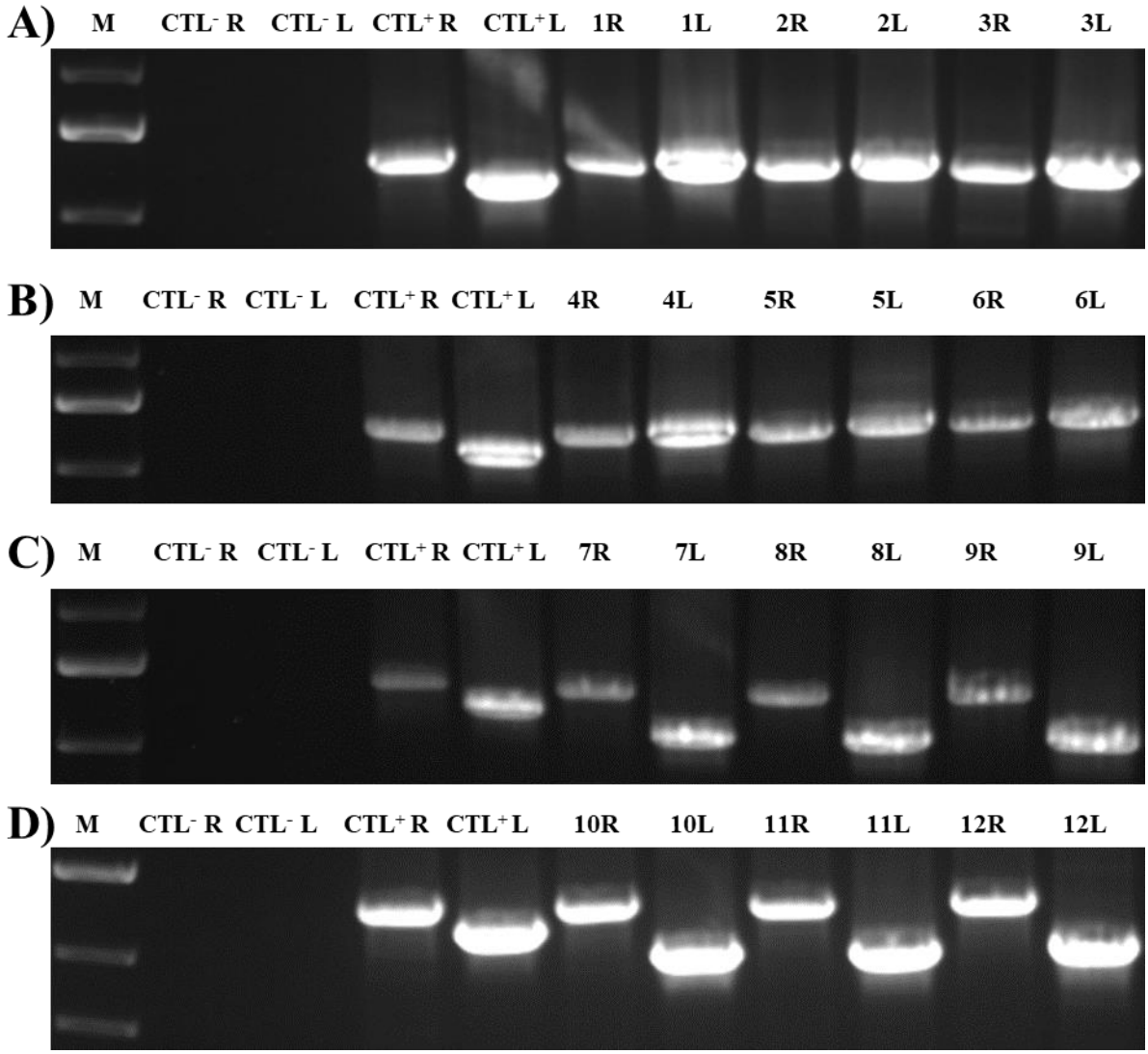
Third experiment



Appendix E Figure 8 (A, B, C, and D): Immunoblotting of PHGDH from koHEK cells transfected transiently with 11 different variants of PHGDH.

Immunoblotting of PHGDH from koHEK cells transfected transiently with 11 different variants of PHGDH; besides koHEK-tWT and empty Vector. KoHEK cells without transfection and wtHEK cells were included also as

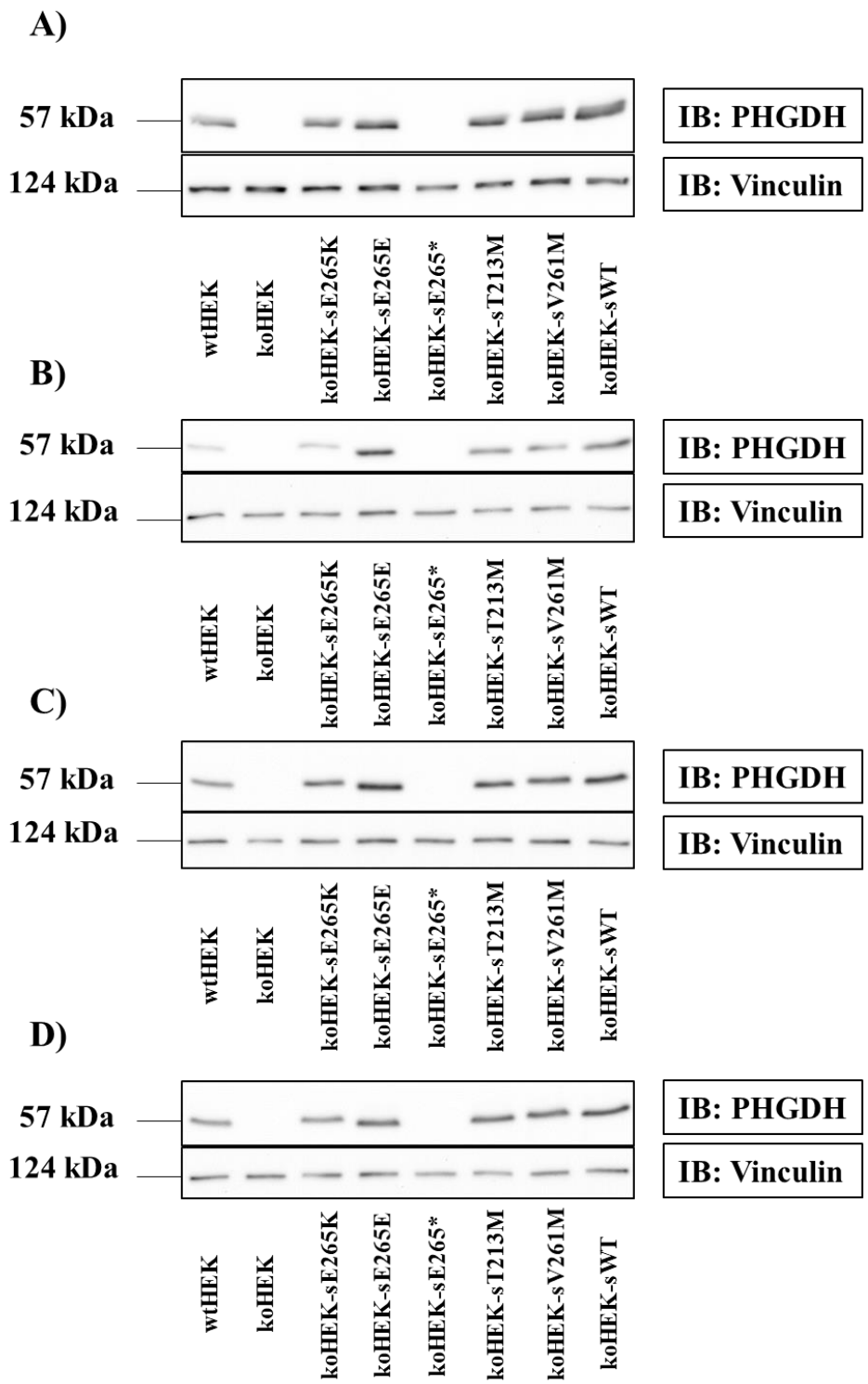
indicated under the figure. Equal loading was confirmed by immunoblotting of Vinculin. Apparent molecular mass markers (in kDa) are indicated and the different molecular mass between PHGDH from wtHEK cells and recombinant proteins, because it is fusion protein to GFP which in turn has a molecular mass of about 27 kDa. Because one Blot was not enough to include all our variants, so we loaded it on 3 different blots, but each blot included wtHEK; koHEK, koHEK-tWT and an empty. Arrows indicate the theoretical molecular mass of the truncated forms of PHGDH.



Appendix E Figure 9 (A, B, Cand D): Junction PCR data demonstrating the functionality of the pZT-AAVS1-TALE Nuclease pair and HDR of donor vectors in koHEK cells to generate knock-in HEK cells.

Genomic DNA samples of knock-in PHGDH HEK cells were analysed for AAVS1 targeted homologous recombination using junction PCR.The presence of the expected band refer to the successful integration. CTL- refers to koHEK cells without transfection with any vector and were used as a negative control. CTL+ refers to koHEK cells co-transfected with pZT-AAVS1-TALE Nuclease pair and positive control donor vector provided with the kit to estimate the functionality of HDR approach. R refers to the primers pair amplify a specific fragment right of AAVS1 locus to validate the integration of the donor vector right of the genome. L refers to the primers pair

amplify a specific fragment left of AAVS1 locus to validate the integration of the donor vector left of the genome. The numbers from 1 to 12 refer to koHEK co-transfected with pZT-AAVS1-TALE Nuclease pair and the twelf constructed donor vectors. ***1*** *refers to* ***koHEK-sE265*-CMV (****koHEK co-transfected with pZT-AAVS1-TALE Nuclease pair and pAAVS1D-CMV-PHGDH (c.793G>T) -EF1a-copGFP-T2A-Puro).* ***2*** *refers to* ***koHEK-sE265K-CMV*** *(koHEK co-transfected with pZT -AAVS1-TALE Nuclease pair and the constructed donor vector pAAVS1D-CMV-PHGDH (c.793G>A) -EF1a-copGFP-T2A-Puro).* ***3*** *refers to* ***koHEK-sE265E-CMV*** *(koHEK co-transfected with pZT -AAVS1-TALE Nuclease pair and the constructed donor vector pAAVS1D-CMV-PHGDH (c.795G>A) -EF1a-copGFP-T2A-Puro).* ***4*** *refers to* ***koHEK-sT213M-CMV*** *(koHEK co-transfected with pZT -AAVS1-TALE Nuclease pair and the constructed donor vector pAAVS1D-CMV-PHGDH (c.638C>T) -EF1a-copGFP-T2A-Puro).* ***5*** *refers to* ***koHEK-sV261M-CMV*** *(koHEK co-transfected with pZT-AAVS1-TALE Nuclease pair and the constructed donor vector pAAVS1D-CMV-PHGDH (c.781G>A) -EF1a-copGFP-T2A-Puro).* ***6*** *refers to* ***koHEK-sWT-CMV*** *(koHEK co-transfected with pZT-AAVS1-TALE Nuclease pair and the constructed donor vector pAAVS1D-CMV-PHGDH (WT)-EF1a-copGFP-T2A-Puro).* ***7*** *refers to* ***koHEK-sE265*-PGK (****koHEK co-transfected with pZT-AAVS1-TALE Nuclease pair and pAAVS1D-PGK-PHGDH (c.793G>T) -EF1a-copGFP-T2A-Puro).* ***8*** *refers to* ***koHEK-sE265K-PGK (****koHEK co-transfected with pZT-AAVS1-TALE Nuclease pair and the constructed donor vector pAAVS1D-PGK-PHGDH (c.793G>A) -EF1a-copGFP-T2A-Puro).* ***9*** *refers to* ***koHEK-sE265E-PGK*** *(koHEK co-transfected with pZT-AAVS1-TALE Nuclease pair and the constructed donor vector pAAVS1D-PGK-PHGDH (c.795G>A) -EF1a-copGFP-T2A-Puro).* ***10*** *refers to* ***koHEK-sT213M-PGK (****koHEK co-transfected with pZT -AAVS1-TALE Nuclease pair and the constructed donor vector pAAVS1D-PGK-PHGDH (c.638C>T) -EF1a-copGFP-T2A-Puro).* ***11*** *refers to* ***koHEK-sV261M-PGK*** *(koHEK co-transfected with pZT -AAVS1-TALE Nuclease pair and the constructed donor vector pAAVS1D-PGK-PHGDH (c.781G>A) -EF1a-copGFP-T2A-Puro).* ***12*** *refers to* ***koHEK-sWT-PGK*** *(koHEK co-transfected with pZT -AAVS1-TALE Nuclease pair and the constructed donor vector pAAVS1D-PGK-PHGDH(WT)-EF1a-copGFP-T2A-Puro). pZT-AAVS1-L1/R1TALE-Nuclease vectors.*



Appendix E Figure 10: (A, B, C, and D): Immunoblotting of PHGDH from knock-in PHGDH HEK cells.

wtHEK and koHEK cells were also included. Equal loading was confirmed by immunoblotting of vinculin. Apparent molecular mass markers (in kDa) are indicated. One could notice almost stable expression of PHGDH in all the stable transfected koHEK cells (All these cells were with CMV promoter).A) shows the results from the first experiment B) shows the results from the second experiment.C) shows the results from the third exxperiment. D) shows the results from the fourth experiment.

Appendix F

Appendix F Table 1: Physical findings and identified variants in my cohort.

General data																			
Patient / family number	1	2	3	4	5a	5b	6a	6b	7a	7b	8	9	10	11	12	13	14	15a	15b
Country of referral	France	Canada	Australia	Iran	Iran	Iran	Turkey	Turkey	Germany	Germany	Germany	Belgium	USA	Ireland	France	Ireland	Spain	New Zealand	New Zealand
Mutated gene	PHGDH	PHGDH	PHGDH	PHGDH	PSAT1	PSAT1	PSAT1	PSAT1	PSAT1	PSAT1	PSAT1	PSAT1	PSAT1	PSAT1	PSAT1	PSAT1	PSAT1	PSAT1	PSAT1
Nucelotide change(s)	c.160C>T(homo)	c.488G>A (homo)	c.638C>T (homo; UPD1)	c.704C>T (homo)	c.1A>G (presumed homo; hetero in parents)	c.1A>G (presumed homo; hetero in parents)	c.129T>G (homo)	Presumably c.129 T>G (homo) (similar phenotype as sibling 6a)	c.181C>T (hetero,mat) c.296C>T (hetero,pat)	c.181C>T (hetero,mat) c.296C>T (hetero,pat)	c.235G>T(homo) (hetero in both parents)	c.296C>T (homo)	c.296C>T(hetero) c.870-1G>T(hetero)	c.296C>T (hetero, mat) **missing 2nd allele**	c.733T>C (presumed homo; hetero in parents)	c.463G>C (hetero) c.870-1G>T (hetero)	c.870-1G>T (presumed homo; hetero in parents)	c.955delA (homo)	c.955delA (homo)
Predicted consenquences on RNA / protein level	p.(Arg54Cys)	p. (Arg163Gln)	p. (Thr213Met)	p.(Ala235Val)	p.(Asp2_Met42del)	p. (Asp2_Met42del)	p.(Ser43Arg)	Presumably, p.(Ser43Arg)	p.(Arg61Trp); p.(Ala99Val)	p.(Arg61Trp); p.(Ala99Val)	p.(Gly79Trp)	p. (Ala99Val)	p.(Ala99Val); Splicing	p. (Ala 99Val)	p.(Cys245Arg)	p. (Glu155Gln); splicing	splicing	p.(Arg319Aspfs*14)	p.(Arg319Aspfs*14)
Mutation screening method	TS	TS	TS	WES	TS	TS	WES	No gene analysis in this patient	WES	TS	WES	TS	TS	TS	TS	TS	TS	TS	TS
Gender	Female	Female	Female	Male	Female	Female	Male	Male	Female	Male	Female	Male	Male	Female	Male	Female	Female	Male	Male
Parental consanguinity (specify)	Yes	None known	None	First cousins	First cousins	First cousins	First cousins	First cousins	None	None	2nd dgree cousins	3rd degree cousins	None	NA	2nd degree cousins	None	None	None known	None known
Affected siblings	None	None	None	Possibly affected fetus in previous pregnancy	One, patient 5b	One, patient 5a	One, patient 6b	One, patient 6a	One, patient 7b	One, patient 7b	Probably affected male sibling	None	None	None	Probably affected male sibling	None	None	One, patient 15b	One, patient 15a
Ethnic background	Caucasian	Pakistani	NA	Caucasian	Caucasian	Caucasian	Turkish	Turkish	Caucasian	Caucasian	Arabian	Arabian	NA	NA	Turkish	NA	Caucasian	Tuvaluans	Tuvaluans
Pregnancy, birth, death																			
Polyhydramnios	No	No	No	Not reported	No	No	No	No	No	NA	No	No	No	No	No	No	No	No	No
Decreased fetal movements	No	Yes	No	Not reported	No	No	No	No	No	NA	Yes	Not reported	Not reported	Yes	Yes	Yes	Not reported	Yes	No
Intrauterine growth restriction (IUGR)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	Yes
Gestational age at birth	26+6 weeks	39+2 weeks	40 weeks	33 weeks	40 weeks	40 weeks	36 w	35 w	32+2 weeks	15 weeks	26 weeks	36+5 weeks	37+2 weeks	39 weeks	32 weeks	38+5 weeks	35 weeks	31+6 weeks #	18+6 weeks
Birth weight	379 g (-2.5 SD)	1988 g (-3.2 SD)	2540 g (-2.0 SD)	1179 g (-2.1 SD)	1900 g (-3.7 SD)	2200 g (-2.9 SD)	2500 g (-1.7 SD)	NA	820 g (-2.4 SD)	33 g	278 g (-3.2 SD)	1330 (-3.8 SD)	1371 g (-4.0 SD)	2120 g (-2.9 SD)	770 g (-2.6 SD)	1507 g (-4.6 SD)	NA	1690 g (-0.2 SD)	96.2 g
Body length at birth	28 cm (-2.7 SD)	42.5 cm (-3.1 SD)	39 cm (-5.0 SD)	31 cm (-4.9 SD)	40 cm (-4.6 SD)	39 cm (-5.0 SD)	47 (-1.1 SD)	NA	33 cm (-3.3 SD)	9 cm	Crown-rump 15.,5 cm	33 cm (-6.2 SD)	38.6 cm (-4.8 SD)	NA	30 cm (-4.7 SD)	Crown-rump 25.3 cm, crown-heel 25.8	NA	43 cm (+0.5 SD)	17.8 cm
OFC at birth	18 cm (-4.5 SD)	27.4 cm (-5.0 SD)	28.2 cm (-4.7 SD)	26 cm (-2.9 SD)	30 cm (-3.4 SD)	29 cm (-4.1 SD)	30 (-2.3 SD)	NA	24 cm (-3.4 SD)	NA		25 cm (-5.5 SD)	24.6 cm (-6.1 SD)	NA	20.5 cm (-6.0 SD)	26.4 cm (-5.4 SD)	NA	29 cm (-0.2 SD)	11.6 cm
Intrauterine fetal death (IUFD) / termination of pre	TOP	No	No	IUFD	No	No	No	No	No	TOP	TOP	No	No	No	TOP but fetal distress announcing IUFD	No	IUFD (stillbirth at 35 weeks of gestation)	No	TOP
Liveborn (indicate age of death)	No	Yes (died on first day)	Yes (died after 30 days)	No	Yes (died after 27 days)	Yes (died after 5 days)	Yes (alive at age 4 years)	Yes (died a age 6.5 years)	Yes (died at age 4 months)	No	No	Yes (died on first day)	Yes (died after 5 days)	Yes (died after 7 days)		Yes (died on first day)	No	Yes (died shortly after birth)	No
Craniofacial																			
Slanted forehead	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	Yes	NA	Yes	Yes	Yes	Yes	Yes	Yes
Short neck	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	NA	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Malformed ears	Yes	Yes	No	Yes	No	No	Yes	Yes	No	NA	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Low-set ears	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	NA	Yes	Yes	Yes	Yes	Mildly	Yes	Yes	No	No
Round gaping mouth	Yes	Yes	No	Yes	No	No	Yes	No	No	NA	No	Yes	No	Yes	Yes	Yes	Yes	No	No
Everted lips	Yes	No	No	Yes	No	No	Yes (lower lip)	No	No	NA	Yes	Yes	Yes	Yes	Yes (lower lip)	Yes	Yes	No	No
Cleft lip and / or cleft palate	No	No	No	No	No	No	Cleft palate	Cleft palate	Cleft palate	No	No	No	No	No	No	Cleft palate	No	No / NA	Cleft palate
Micro- / retrognathia	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other craniofacial abnormalities (please specify)	Short nose, anteverted nostrils	Hypoplastic malar areas, broad nose, bilateral epicanthic folds, hypertelorism					Epicanthus inversus, prominent glabella,	Flattened nasal root, epicanthus inversus				Hypertelorism			Marked hypertelorism, flattened nose, hypoplastic nares, long philtrum, low-set cheeks		Facial edema, hypertelorism, downslanting palpebral fissures, small nose, anteverted nares		
Ocular																			
Proptosis	No	Yes	Yes	Yes	NA	NA	Yes	Yes	No	NA	NA	Yes	No	NA	No	Yes	Yes, mild	No	No
Shortened / everted eyelids	Yes	No	No	Yes	Yes	Yes	Yes	NA	No	NA	NA	Yes	No	Yes	No	Yes	Yes	No	No
Cataracts	Yes (bilateral)	Yes	No	No	Yes	Yes	Yes	Yes	No	NA	NA	No	Yes	No	Yes	No	NA	NA	NA
Other ocular abnormalities (please specify)	Irregular fibers of crystalline with central calcification						Vision problems due to cataract	Vision problems due to cataract					One eyelid fused; at post mortem exam: persistent fetal vasculature retinal detachment, optic nerve hypoplasia		Microcornea	Absence of eyelids			
Central nervous system																			
Microcephaly	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lissencephaly / abnormal gyration (specify)	Yes	Yes	Yes, some lissencephaly	Yes	NA	NA	Deep cortical sulcus	NA	No	NA	Yes, pachygyria	Yes	Yes	NA	Yes, simplified gyration with disorganisation of the cortex	yes	Not on antental ultrasound; post mortem not done	Not on antental ultrasound; post mortem not done	NA
Cerebellar hypoplasia	Yes	Yes	Yes	No	NA	NA	No	NA	No	NA	aplasia vermis cerebelli	No	Yes	NA	Yes (no identifiable cerebellum)	Yes	Not on antental ultrasound; post mortem not done	Yes	Abnormal cerebellum on ultrasound prior to TOP
Corpus callosum hypoplasia / agenesis	Yes	Yes	Yes, agenesis	Yes	NA	NA	Yes	NA	Yes, agenesis	NA	Yes, agenesis	Yes, agenesis	Yes, agenesis	NA	Yes, agenesis	No	Not on antental ultrasound; post mortem not done	Not on antental ultrasound; post mortem not done	Not on antenatal ultrasound
Hydrocephalus / enlarged ventricles	yes (moderate)	enlarged ventricles	yes	no	NA	NA	No	NA	yes	NA	yes	enlarged ventricles	yes	NA	no	yes	Not on antental ultrasound; post mortem not done	Not on antental ultrasound; post mortem not done	Not on antenatal ultrasound
Intracerebral calcifications	Yes	Yes	unknown	No	NA	NA	No	NA	No	NA	No	No	No	NA	no	no	Not on antental ultrasound; post mortem not done	Not on antental ultrasound; post mortem not done	NA

Spina bifida	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Other CNS abnormalities (please specify)		Diffuse pachygyria with thinned, abnormally laminated neocortex; diffuse severe telencephalic white matter rarefaction/cavitation predominantly involving the subplate with periventricular cysts; widespread gliosis and aplasia of the corticospinal tracts; thalamus and basal ganglia small and gliotic: cerebellar hypoplasia with atrophy of the folia and depletion of Purkinje cell, external and internal granular layers in the neocerebellum	Sub-ependymal pseudocyst				Linear signal increase in FLAIR images on ependymal and sulcal faces					Absence of septum pellucidum	Reduced number of anterior horn motor neurons at autopsy		Large meningeal capillary vessels, large vessels in the germinative zone which is sparse, dysplastic thalami			Dandy-Walker malformation	
CNS findings verified by imaging /autopsy	Autopsy	Autopsy		Autopsy	None	None	Brain MRI	None					Autopsy				Antenatal ultrasound	Antenatal ultrasound	Antenatal ultrasound
Limbs / skeltal																				
Joint contractures	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Swelling of hands	No	Yes	Yes	Yes	No	No	No	No	No	NA	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes	No
Finger syndactyly	No	No	No	Yes	No	No	No	No	No	NA		Yes	No	Yes	No	Yes	Not described, but the fingers are just minimal appendages in the edematous hands	No	No	
Swelling of feet	No	Yes	No	Yes	No	No	Yes	NA	No	NA		Yes	NA	Yes	No	Yes	Yes	Yes	No	
Toe syndactyly	No	No	No	Yes	No	No	No	No	No	NA		Yes	No	Yes	No	Yes	Not described, but the toes, like the fingers, are just minimal appendages in the edematous hands	No	No	
Rockerbottom feet	No	Yes	Yes	Yes	No	No	Yes	NA	No	NA	Yes	Yes	Yes	Yes	Yes	Yes (right)	Extreme malposition of feet	Yes	Yes	
Clubfeet	Pes varus	No	No	No	No	No	No	NA	No	NA		Yes	Yes	Yes	No	Yes (left)	Extreme malposition of feet	No	No	
Other skeletal abnormalities (please specify)	Distal arthrogryposis, adducted thumbs, thin overlapping fingers	Prominent heels, transverse palamar creases		Scoliosis			Clenched fists	Clenched fists	Preaxial hexadactyly on the right side	Pectus excavatum	Clinodactyly on both sides, hypoplastic thumbs	Sacral agenesis	Clenched fists		Short tubular bones, narrow thorax, small hole on the occipital bone, mild hypomineralisation of the skeleton		Swelling of all joints, especially the elbows, wrists, knees and ankles; short fingers and toes	Hypomineralised long bones on antenatal ultrasound		
Additional																				
Ichthyosis	Yes	Yes	Scaling of skin	Yes	Yes	Yes	Yes, very mild	NA	No	NA	No	Yes	Yes	Yes	Yes	Yes	No	No	No	
Restrictive dermopathy	No	Yes	No	Yes	No	No	No	No	No	NA	No	NA	Yes	Yes	No	Yes	Yes	Yes	Tight facial skin	
Short umbilical cord	Yes, 18 cm	Yes	No	NA	No	No	NA	NA	No	NA	No	No	NA	NA	Not reported	Yes, 19 cm	NA	Not reported	Not reported	
Heart defect (specify)	No	No	No	No	No	No	No	No	No	No	No	No	No	NA	Yes, atrioventricular septal defect, dextrocardia, hypoplastic right ventricle and pulmonary artery	Yes, biventricular hypertrophy	Not on antental ultrasound; post mortem not done	Not on antental ultrasound; post mortem not done	No	
Hypoplastic lungs	No	Yes, with incomplete minor fissure of right lung and major fissure of left lung	No	Yes	NA	NA	No	No	No	NA	No	Not reported	Yes	Yes	No	Yes	NA	not on antental ultrasound; post mortem not done	No	
Endocrine / metabolic abnormalities (specify)	No	No	No	No	NA	NA	His: 27,9 (41-125), Ser: 47,16 (69-187), Anserin: 0 (31-51), Isoleucin: 0 (33-97)	NA	No	NA	No	NA	No	NA	NA	Normal adrenal and thyroid	NA	NA	NA	
Renal abnormalities (specify)	No	No	Small kidneys: R 3.8cm, L 3.4cm <5th centile	NA	NA	NA	No	No	No	No	No	Enlarged bladder, pyelic dilation, megaureteres	No	NA	Small kidneys (weight corresponding to 24 weeks gestation)	No	Not on antental ultrasound; post mortem not done	One kidney small with multiple cysts; other kidney large with pyelectasis	Bilateral hydronephrosis on antenatal scan but not on postmortem	
Hypoplastic genitalia	Yes	External genitalia posteriorly displaced, small ovaries	No	Yes	No	No	Yes	NA	Yes	NA	No	Yes	Yes	No	No	No	No	Yes	No	
Ambiguous genitalia	No	No	No	No	No	No	No	NA	No	NA	No	Micropenis	No	No	No	Everted vagina and labia	Yes	Yes	No	
Other (please specify)	Wide anus, amyotrophy						No speech, hirsuitism, inguinal hernia, seizures, feeding diffuculty, dystonic posture; last measurements (age 4): OFC 42.5 cm (-5.2 SD), weight 102 cm (-0.5 SD), length 15.5 kg (-0.6 SD)		Duodenal atresia, annular pancreas	Gut malformation,	Prevertebral and cephalic edema of the skin	Hydrops (ascites, prefrontal edema and pleural effusion)			Total situs inversus, posteriorly-set anus, sacral dimple, severe hypoplasia of all skeletal muscles (including facial muscles) which are replaced by fat tissue	Everted anus	Generalized oedema; amyotrophy of limb muscles, more marked distally	Single umbilical artery		

Intrauterine fetal death (IUFD); not availabe / not assessed (NA); occipito-frontal head circumference (OFC); standard deviation (SD); termination of pregnancy (TOP).

Supporting information regarding Appendix F Table 2, 3, and 4

The following three tables provide an overview on reported variants in NLS and SBDNL. Novel variants that are first reported in a scientific article here are printed in bold and all variants observed in the present cohort (comprising 15 unrelated families) are highlighted in green.

Interpretation of variants as NLS-associated or SBDNL-associated change was primarily based on the reported phenotype in homozygous individuals. For variants that have not been observed in a homozygous constellation, thus far, the interpretation was as follows: (i) Frameshift variants and one multi-exon deletion of *PHGDH*, which have only been observed in compound heterozygosity with a non-truncating variant and not in the homozygous state were classified as probable LoF (loss-of-function) alleles. (ii) Variants that have been observed in an individual with SBDNL in compound heterozygosity with a variant classified as NLS-associated variant or probable LoF allele were classified as probable SBDNL-associated variant, because it was evident that this allele confers the less severe phenotype (presumably higher residual enzyme activity). (iii) Non-truncating variants that have been observed in an individual with NLS in compound heterozygosity with a variant classified as NLS-associated variant or probable LoF allele were classified as probable NLS-associated variant, because it was evident that this allele was not able to ameliorate the phenotype towards a less severe non-lethal condition. (iv) Two non-truncating (e.g., missense), not otherwise classified variants observed in compound heterozygosity in association with a NLS phenotype were both classified as possible NLS-associated variants, because it may be assumed that neither of the two variants was able to ameliorate the phenotype towards a less severe non-lethal condition.

Variants refer to the reference sequences of *PHGDH* (NM_006623.4; NG_009188.1), *PSAT1* (NM_058179.4; NG_012165.1) and PSPH (NM_004577.4; NG_011473.1). Compliance with HGVS nomenclature was verified using the Mutalyzer program (https://mutalyzer.nl/).

Further legends are provided for each table separately.

Appendix F Table 2: PHGDH variants and associated phenotypes.

Variant	Consequence on RNA /protein	Homozygous phenotype	Reference	Comp. het. phenotype	Second allele(s)	Reference	Interpretation
c.1A>C	p.?	NLS	(Bourque *et al.*, 2019)				NLS-associated variant
c.138+2dup	p.? ¶	NO		SBDNL	c.1129G>A	(Benke *et al.*, 2017; Glinton *et al.*, 2018)	Not classified
c.160C>T	p.(Arg54Cys)	NLS	(Abdelfattah *et al.*, 2020)	NLS	Exon del	(Acuna-Hidalgo *et al.*, 2014)	NLS-associated variant
c.164C>T	p.(Ser55Phe)	NO		SBDNL	c.1429dup	(Poli *et al.*, 2017)	Probable SBDNL-associated variant
c.277A>T	p.(IIe93Phe)	NLS	(Kapoor *et al.*, 2021)	NO			NLS-associated variant
c.403C>T	p.(Arg135Trp)	NO		SBDNL	c.487C>T	(Brassier *et al.*, 2016)	Probable SBDNL-associated variant
					c.712delG	(Tabatabaie *et al.*, 2009)	
c.418G>A	p.(Gly140Arg)	NLS	(El-Hattab *et al.*, 2016; Shaheen *et al.*, 2014)	NO			NLS-associated variant
c.487C>T	p.(Arg163Trp)	NO		SBDNL	c.403C>T	(Brassier *et al.*, 2016)	Not classified
c.488G>A	p.(Arg163Gln)	NLS	(Abdelfattah *et al.*, 2020; Shaheen *et al.*, 2014)	NO			NLS-associated variant

c.638C>T	**p.(Thr213Met)**	NLS	(Abdelfattah *et al.*, 2020)	NO			NLS-associated variant
c.704C>T	**p.(Ala235Val)**	NLS	(Abdelfattah *et al.*, 2020)	NO			NLS-associated variant
c.714del †	p.(Ile239Serfs*69)	NO		SBDNL	c.403C>T	(Tabatabaie *et al.*, 2009)	Probable LoF allele
c.746T>C	p.(Leu249Pro)	NLS	(Cavole *et al.*, 2020)	NO			NLS-associated variant
c.781G>A	p.(Val261Met)	SBDNL	(Coşkun *et al.*, 2009; Tabatabaie *et al.*, 2009)	NO			SBDNL-associated variant
c.793G>A	p.(Glu265Lys)	NLS	(Acuna-Hidalgo *et al.*, 2014)	NO			NLS-associated variant
c.856G>C	p.(Ala286Pro)	NLS	(Acuna-Hidalgo *et al.*, 2014)	NO			NLS-associated variant
c.1030C>T	p.(Arg344*)	NIHF ‡	(Monies *et al.*, 2017)	NO			Probable LoF allele
c.1117G>A	p.(Ala373Thr)	SBDNL	(Tabatabaie *et al.*, 2009)	NO			SBDNL-associated variant
c.1129G>A	p.(Gly377Ser)	SBDNL	(Glinton *et al.*, 2018; Tabatabaie *et al.*, 2009)	SBDNL	c.138+2dup	(Benke *et al.*, 2017)	SBDNL-associated variant
c.1219T>C	p.(Ser407Pro)	SBDNL	(Ali *et al.*, 2021)	NO			SBDNL-associated variant
c.1263C>G	p.(Cys421Trp)	NO		NLS	c.1468G>A	(Ni *et al.*, 2019)	Probable NLS-associated variant

c.1273G>A	p.(Val425Met)	SBDNL	(Klomp *et al.*, 2000; Kraoua *et al.*, 2013)	SBDNL	c.1471C>T	(Meneret *et al.*, 2012)	SBDNL-associated variant
c.1286G>T	p.(Gly429Val)	SBDNL/NLS	(Benke *et al.*, 2017; El-Hattab *et al.*, 2016)	NO			Intermediate variant
c.1297C>T	p.(Gln433*)	NLS	(Mattos *et al.*, 2015)	NO			NLS-associated variant
c.1429dup	p.(Met477Asnfs*51)	NO		SBDNL	c.164 C>T	(Poli *et al.*, 2017)	Probable LoF allele
c.1468G>A	p.(Val490Met)	SBDNL	(Klomp *et al.*, 2000; Pind *et al.*, 2002)	NLS	c.1263C>G	(Ni *et al.*, 2019)	SBDNL-associated variant / Intermediate variant
c.1471C>T	p.(Arg491Trp)	NO		SBDNL	c.1273G>A	(Meneret *et al.*, 2012)	Not classified
c.1518 G>A	p.Trp506*	NO		NLS	Maternal inversion of 1,2 Mb bases affects PHGDH	(Takeichi *et al.*, 2018)	Probable LoF allele
Del. exons 10-12	p.?	NO		NLS	c.160C>T	(Acuna-Hidalgo *et al.*, 2014)	Probable LoF allele

Compound heterozygous (Comp. het); deletion (Del); loss-of-function (LoF); Phenotype: non-immune hydrops fetalis (NIHF); Neu-Laxova syndrome (NLS); not observed / not reported (NO); serine biosynthesis defect, non-lethal (SBDNL).
† Reported in the literature as c.712delG. ¶ Variant is assumed to affect splicing.
‡ Only rudimentary clinical information provided for this case reported with non-immune hydrops fetalis.
Variant observed in the present cohort. ***Novel variant***

Appendix F Table 3: PSAT1 variants and associated phenotypes.

Variant	Consequence on RNA /protein	Homozygous phenotype	Reference	Comp. het. phenotype	Second allele(s)	Reference	Interpretation
c.1A>G	**p.?**	NLS	(Abdelfattah *et al.*, 2020)	NO			NLS-associated variant
c.43G>C	**p.(Ala15Pro)**	SBDNL	(Shen *et al.*, 2022)	SBDNL	c.467C>T	(Debs *et al.*, 2021)	SBDNL-associated variant
c.44C>T	p.(Ala15Val)	NO		SBDNL	c.432delA	(Glinton *et al.*, 2018)	Probable SBDNL-associated variant
c.107del	p.(Gly36Alafs*7)	NO		SBDNL	c.299A>C	(Hart *et al.*, 2007)	Probable LoF allele
c.129T>G	p.(Ser43Arg)	SBDNL	(Abdelfattah *et al.*, 2020; Brassier *et al.*, 2016)	NO			SBDNL-associated variant
c.181C>T	**p.(Arg61Trp)**	NO		NLS	c.296C>T	(Abdelfattah *et al.*, 2020)	Probable NLS-associated variant
c.208T>A	p.(Tyr70Asn)	NO		NLS	c.1024C>T	(Ni *et al.*, 2019)	Possible NLS-associated variant
c.233G>C	p.(Gly78Ala)	MCA ‡	(Monies *et al.*, 2017)	NO			Not classified ‡
c.235G>T	**p.(Gly79Trp)**	NLS	(Abdelfattah *et al.*, 2020)	NO			NLS-associated variant

c.296C>T	p.(Ala99Val)	NLS	(Abdelfattah *et al.*, 2020; Acuna-Hidalgo *et al.*, 2014; El-Hattab *et al.*, 2016)	NLS	c.536C>T	(Acuna-Hidalgo *et al.*, 2014)	NLS-associated variant (common allele)
					c.181C>T	(Abdelfattah *et al.*, 2020)	
					c.870-1G>T	(Abdelfattah *et al.*, 2020)	
					Missing allele	(Abdelfattah *et al.*, 2020)	
c.299A>C	p.(Asp100Ala)	NO		SBDNL	c.107delG	(Hart *et al.*, 2007)	SBDNL-associated variant
c.432del	p.(Asp145Metfs*49)	NO		SBDNL	c.44C>T	(Glinton *et al.*, 2018)	Probable LoF allele
c.463G>C	**p.(Glu155Gln)**	NO		NLS	c.870-1G>T	(Abdelfattah *et al.*, 2020)	Probable NLS-associated variant
c.467C>T	**p.Thr156Met**	NO		SBDNL	c.43G>C	(Debs *et al.*, 2021)	Probable SBDNL-associated variant
c.536C>T	p.(Ser179Leu)	NLS	(Acuna-Hidalgo *et al.*, 2014)	NLS	c.296C>T	(Acuna-Hidalgo *et al.*, 2014)	NLS-associated variant
c.733T>C	**p.(Cys245Arg)**	NLS	(Abdelfattah *et al.*, 2020)	NO			NLS-associated variant
c.870-1G>T	**p.?** ¶	NLS	(Abdelfattah *et al.*, 2020)	NLS	c.296C>T	(Abdelfattah *et al.*, 2020)	NLS-associated variant

					c.463G>C	(Abdelfattah *et al.*, 2020)	
c.955del	**p.(Arg319Aspfs*14)**	NLS	(Abdelfattah *et al.*, 2020)	NO			NLS-associated variant
c.1023_1027 delinsAGACCT	p.(Arg342Aspfs*6)	NLS	(Acuna-Hidalgo *et al.*, 2014)	NO			NLS-associated variant
c.1024C>T	p.(Arg342Trp)	NO		NLS	c.208T>A	(Ni *et al.*, 2019)	Possible NLS-associated variant

Compound heterozygous (Comp. het); loss-of-function (LoF); Phenotype: multiple congenital anomalies (MCA); Neu-Laxova syndrome (NLS); not observed / not reported (NO); serine biosynthesis defect, non-lethal (SBDNL)
‡ Only rudimentary clinical information provided for this case; therefore, no interpretation is provided. Variant is assumed to affect splicing.
Variant observed in the present cohort. ***Novel variant***

Appendix F Table 4: PSPH variants and associated phenotypes.

Variant	Consequence on RNA /protein	Homozygous phenotype	Reference	Comp. het. phenotype	Second allele(s)	Reference	Interpretation
c.94G>A	p.(Asp32Asn)	NO		SBDNL	c.155T>C	(Veiga-da-Cunha *et al.*, 2004)	Not classified
c.103G>A	p.(Ala35Thr)	SBDNL	(Vincent *et al.*, 2015)	SBDNL	NO		SBDNL-associated variant
c.131T>G	p.(Val44Gly)	NO		SBDNL	c.421G>A	(Byers *et al.*, 2016)	Not classified
c.155T>C	p.(Met52Thr)	NO		SBDNL	c.94G>A	(Veiga-da-Cunha *et al.*, 2004)	Not classified
c.267del	p.(Gly90Alafs*2)	NLS	(Acuna-Hidalgo *et al.*, 2014)	NO			NLS-associated variant
c.421G>A	p.(Gly141Ser)	NO		SBDNL	c.131T>G	(Byers *et al.*, 2016)	Not classified

Compound heterozygous (Comp. het); Phenotype: Neu-Laxova syndrome (NLS); not observed / not reported (NO); serine biosynthesis defect, non-lethal (SBDNL).