Epigenetic regulation of *PLS3* by the macrosatellite *DXZ4* and the transcriptional regulator CHD4

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Eike Andreas Strathmann

aus Kiel

Köln, 2022

The Doctoral Thesis “Epigenetic regulation of *PLS3* by the macrosatellite *DXZ4* and the transcriptional regulator CHD4” was performed at the Institute of Human Genetics and Center for Molecular Medicine Cologne (CMMC) of the University of Cologne from March 2016 to August 2022. The project was performed by Eike A. Strathmann and was supervised by Prof. Dr. Brunhilde Wirth (first supervisor), Prof. Dr. Michael Nothnagel (additional supervisor) and Prof. Dr. Martin Peifer (additional supervisor).

Berichterstatter:	Prof. Dr. Brunhilde Wirth
	Prof. Dr. Niels Gehring
Tag der mündlichen Prüfung:	14.11.2022

Life has a melody, Gaius. A rhythm of notes that become your existence once played in harmony with God's plan. It's time to do your part and realize your destiny.

- Caprica 6

Für meine Familie, die immer an mich geglaubt hat.

Table of content

Contents

Abbreviations

General abbreviations

AAV	adeno associated virus
ASO	antisense oligonucleotide
BWA	Burrows-Wheeler Alignment
CAKUT	Congenital Anomalies of the Kidney and Urinary Tract
cDNA	complementary DNA
ChIP-qPCR	chromatin immunoprecipitation assays followed by qRT-PCR
CHOP-INTEND	Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders
CTC	circulating tumour cells
EBV	Epstein-Barr Virus
EMT	epithelial-mesenchymal transition
FB	fibroblast
gDNA	genomic DNA
iPSC	induced pluripotent stemm cell
KW test	Kruskal Wallis test
lncRNA	long non-coding RNAs
MHFMS	Modified Hammersmith Functional Motor Scale
miRNA	microRNA
MN	(spinal) motor neuron
mRNP	mRNA-binding proteins
NMJ	neuromuscular junction
NuRD	Nucleosome Remodelling Complex
PAR1	pseudoautosomal region 1
PAR2	pseudoautosomal region 2
pNF-H	phosphorylated heavy-chain neurofilaments
RPMI	Reads per Kilo Base per Million Mapped Reads
SMA	Spinal Muscular Atrophy
SMA1	SMA type 1, severe form
SMA2	SMA type 2, intermediate form
SMA3	SMA type 3, mild form
SMA4	SMA type 4, adult onset
snRNP	small nuclear ribonucleoprotein
SNV	single nucleotide variant
SS	Sézary Syndrome
TAD	topologically associating domain
TIMPSI	Test for Infant Motor Performance Items
TSS	transcription start sites
WGS	whole-genome sequencing
WSR test	Wilcoxon-Signed-Rank test
X_a	active X-chromsome
XCI	X-chromosomal inactivation
X_i	inactivated X-chromosome
XI status	X-inactivation status

Gene and protein names

AHSG	Alpha 2-HS Glycoprotein, Fetuin A
CHD3	Chromodomain Helicase DNA Binding Protein 3
CHD4	Chromodomain Helicase DNA Binding Protein 4
CHD5	Chromodomain Helicase DNA Binding Protein 5
CLEC3B	C-type Lectin Family 3 Member B, Tetranectin
CMV	Cytomegalovirus
COMP	Cartilage Oligomeric Matrix Protein
CTCF	CCCTC-binding factor
CTCF	CCCTC-Binding Factor
DANT1	DXZ4-Associated Non-Coding Transcript 1
DANT2	DXZ4-Associated Non-Coding Transcript 2
DPP4	Dipeptidyl Peptidase 4
DXZ4	DXZ4 macrosatellite
GATA2	GATA Zinc Finger Domain Containing Potein
GATA6	GATA Binding Protein 6
H19	Imprinted maternally expressed noncoding transcript
HDAC1	Histone Deacetylase 1
HDAC2	Histone Deacetylase 2
HPRT	Hypoxanthine Phosphoribosyltransferase 1
IgG	Immunoglobulin Heavy Constant Gamma 1
LCP1	Lymphocyte Cytosolic Protein 1, l-Plastin, Plastin 2
MB	Myoglobin
MBD2	Methyl-CpG Binding Domain Protein 2
MBD3	Methyl-CpG Binding Domain Protein 3
MORF4L2	Mortality Factor 5 Like 2
MTA1	Metastasis Associated 1
MTA2	Metastasis Associated 2
MTA3	MetastasisAssociated 3
NCALD	Neurocalcin Delta
PLS1	Plastin 1, i-Plastin
PLS3	Plastin 3, t-Plastin
RBBP4	Retinoblastoma Binding Protein 4
RBBP7	Retinoblastoma Binding Protein 7
SAC6	Fimbrin
SLITRK4	SLIT And NTRK Like Family Member 4
SMN1	Survival of Motor Neuron 1
SMN2	Survival of Motor Neuron 2
SPP1	Secreted Phosphoprotein 1, Osteopontin
TCEAL4	Transcription Elongation Factor A Like 4
TWIST1	Twist Family BHLH Transcription Factor 1
VTN	Vitronectin
XIST	X-Inactive Specific Transcript
ZFHX4	Zink Finger Homeobox 4

1. Declaration of contributions

1.1. Publications

Parts of the data, tables, paragraphs and figures of this doctoral thesis have been adapted from the following original research articles and reviews, which have been published or are under review:

2022: Scientific publication as author (under review)
Strathmann EA, Hölker I, Tschernoster N, Hosseinibarkooie S, Come J, Martinat C, Altmüller J, Brunhilde Wirth B
Epigenetic regulation of plastin 3 expression by the macrosatellite DXZ4 and the transcriptional regulator CHD4
Under revision at AJHG

2021: Review as shared author
Wolff L, **Strathmann EA**, Müller I, Mählich D, Veltmann C, Niehoff A, Wirth B
Plastin 3 in health and disease: a matter of balance
Cell Mol Life Sci. 2021 Jul;78(13):5275-5301.doi: 10.1007/s00018-021-03843-5. Epub 2021 May 23.

2018: Scientific publication as author
Strathmann EA, Peters M, Hosseinibarkooie S, Rigo FW, Bennett CF, Zaworski PG, Chen KS, Nothnagel M, Wirth B
Evaluation of potential effects of Plastin 3 overexpression and low-dose SMN-antisense oligonucleotides on putative biomarkers in spinal muscular atrophy mice.
PLoS One. 2018 Sep 6;13(9):e0203398. doi: 10.1371/journal.pone.0203398. eCollection 2018.

1.2. Methods

Eike A. Strathmann designed and statistically analysed all experiments under supervision of Prof. Dr. Brunhilde Wirth. Contributions to the laboratory work are listed here.

1.2.1. Part I: Expression of PLS3 in health and disease

Method	Contributors
Establishment of EBV and fibroblast cell lines	Various (former) members of the work group of Prof. Dr. Brunhilde Wirth
Establishment of iPSC-derived spinal motor neurons and RNA extraction for sequencing	Dr. Ludwig Heesen Dr. Cecile Martinat (Heesen et al., 2010; Heesen et al., 2016)
Whole-genome and transcriptome sequencing	DECODE, Iceland.
Analysis of transcriptome data from iPSC-derived spinal motor neurons, EBV cells and fibroblasts	Eike A Strathmann
Maintenance of EBV and fibroblast cell lines Extraction of RNA for sequencing	Eike A Strathmann Irmgard Hölker
Whole-genome sequencing data sets	Previously established by members of the work group of Prof. Dr. Brunhilde Wirth
Molecular Combing	Eike A Strathmann (laboratory work, analysis of data, statistics) Irmgard Hölker (laboratory work) Nikolai Tschernoster (laboratory work) Dr. Janine Altmüller (supervision)
qRT-PCR Sanger sequencing (conducted by GATC Biotech AG)	Eike A Strathmann (laboratory work, analysis, figure design) Irmgard Hölker (laboratory work)
Methylation sensitive X-inactivation assay	Eike A Strathmann (laboratory work, analysis)
siRNA-mediated knock-down of CHD4 Overexpression of CHD4	Eike A Strathmann (laboratory work, analysis)
ChIP-qRT-PCR	Eike A Strathmann (Analysis of data, figure design) Irmgard Hölker (laboratory work)
Dual-luciferase promoter assay	Eike A Strathmann (Analysis of data, figure design)
Statistical analysis and writing of the manuscript	Eike A Strathmann supervised by Prof. Dr. Brunhilde Wirth

1.2.2. Part II: The impact of PLS3 on biomarkers of SMA

Method	**Contributors**
Study design	Prof. Dr. Brunhilde Wirth Dr. Miriam Peters
Maintenance of mice Sampling of blood	Dr. Miriam Peters
ECL immunoassay/ analysis of plasma concentrations of biomarkers	PharmOptima
Statistical analysis and writing of the manuscript	Eike A. Strathmann supervised by Prof. Dr. Brunhilde Wirth

1.3. Figures

All figures have been generated by Eike A. Strathmann under supervision of Prof. Dr. Brunhilde Wirth.

1.4. Tables

All tables have been generated by Eike A. Strathmann under supervision of Prof. Dr. Brunhilde Wirth.

2. Summary

Part I: Spinal muscular atrophy (SMA) is a devastating motor neuron disorder caused by homozygous loss of the Survival of Motor Neuron 1 (*SMN1*) gene and insufficient functional SMN protein produced by the *SMN2* copy gene. Overexpression of Plastin 3 (PLS3) fully protects from mild types of SMA. The mechanisms that regulate *PLS3* expression are not fully understood. We performed a multi-omics analysis of SMA-discordant families using lymphoblastoids, fibroblasts and iPSC-derived spinal motor neurons and identified mechanisms, which contribute to the regulation of *PLS3*. We found a 1-fold expression difference in spinal motor neurons of *PLS3* between female asymptomatic and their SMA-affected brothers, which can be explained by the gene's escape from X-chromosomal inactivation. The X-linked *PLS3* is located in close proximity to *DXZ4*, a microsatellite, which is essential for X-chromosomal inactivation. By Molecular Combing, we measured the copy number of *DXZ4* and found a significant correlation with the expression of *PLS3* in females. Additionally, we identified Chromodomain Helicase DNA Binding Protein 4 (CHD4) as an epigenetic transcriptional regulator of *PLS3*. By application of siRNA-mediated knock-down, overexpression, chromatin immunoprecipitation and promoter luciferase assays, we validated the regulation of *PLS3* by CHD4. Thus, we provide evidence for a multilevel epigenetic regulation of *PLS3*.

Part II: Spinraza, a *SMN* antisense oligonucleotide (ASO) that restores full-length *SMN2* transcripts, has been FDA- and EMA-approved for SMA therapy. Hence, the availability of biomarkers allowing reliable disease monitoring would be of great importance. The BforSMA study identified about 200 putative SMA biomarkers. We took advantage of a previously developed intermediate SMA mouse model, treated with presymptomatic low-dose *SMN*-ASO injections and measured the plasma concentrations of SMN and six other SMA biomarkers from the BforSMA study, namely Cartilage Oligomeric Matrix Protein (COMP), Dipeptidyl Peptidase 4 (DPP4), Tetranectin (C-type Lectin Family 3 Member B, CLEC3B), Osteopontin (Secreted Phosphoprotein 1, SPP1), Vitronectin (VTN) and Fetuin A (Alpha 2-HS Glycoprotein, AHSG). Only COMP and DPP4 showed high and SPP1 moderate correlation with the SMA phenotype. *PLS3* overexpression from a human transgene neither influenced the SMN level nor the six biomarkers, supporting the hypothesis that PLS3 acts as an independent protective modifier of SMA.

3. Introduction

3.1. Part I: Expression of PLS3 in health and disease

3.1.1. PLS3 – a versatile structural protein

The main function of Plastin 3 (PLS3, [MIM: 30013]) is binding and bundling of F-actin, a main component of the cytoskeleton, which is formed from G-actin monomers (Bretscher, 1981; Karpova et al., 1995; Pollard and Cooper, 2009; Schwebach et al., 2020; Shinomiya, 2012). The protein is involved in many important F-actin dependent processes, such as cell motility, focal adhesion, cell division, endo- and exocytosis, neurotransmission, vesicle trafficking, axonal local translation, and pathogen infection, among others (Wolff et al., 2021).

PLS3 belongs to a gene family consisting of three tissue-specific paralogs with a sequence identity of about 70% (Shinomiya, 2012). Plastin 1 (*PLS1,* [MIM: 602734]), formerly known as i-Plastin, is located on chromosome 3 (Chr3: 142,315,235-142,432,506 (GRCh37/hg19)). The gene is expressed in the tissues of the inner organs, such as intestine and colon, the kidney and the hair cell stereocilia of the inner ear and is involved in different types of deafness (Lin et al., 1999; Shinomiya, 2012). The second isoform is called lymphocyte cytosolic protein 1 (*LCP1*, [MIM: 153430]), formerly known as l-Plastin or Plastin 2 and is located on chromosome 13 (Chr13: 46,700,058-46,756,312 (GRCh37/hg19)). *LCP1* is expressed primarily in the hematopoietic system and is associated with several malignancies (Samstag and Klemke, 2007; Zeng et al., 2021). PLS3 (or t-Plastin) was first described in 1980 by Anthony Bretcher and Klaus Weber, which named the protein fimbrin (Bretscher and Weber, 1980). The yeast homolog, *SAC6* is still known as Fimbrin. In humans and mice, *PLS3*/*Pls3* is X-linked (humans: ChrX: 114,795,501-114,885,181 (GRCh37/hg19), mice: ChrX: 75,785,654-75,875,182 (GRCm39)).

The transcriptional regulation of *PLS3* in a tissue-specific manner seems to be inevitable for the maintenance of several cellular functions, as gene knock-down or mutations are associated with Osteoporosis, while overexpression can lead to osteoarthritis (Mählich et al., 2021; Tsolis et al., 2015; van Dijk et al., 2013). Increased *PLS3* levels are a hallmark of several malignancies of solid and hematopoietic cells (Kujawski et al., 2015; Kurashige et al., 2019; Kuriyama et al., 2021; Sugimachi et al.,

2014; Ueo et al., 2015; Xin et al., 2020; Yokobori et al., 2013). Most importantly, PLS3 has been identified as a rescuing genetic modifier of Spinal Muscular Atrophy (SMA, [MIM: 253300]) (Heesen et al., 2016; Oprea et al., 2008). Nevertheless, the mechanisms that regulate the expression of PLS3 in various tissues are still enigmatic and several mechanisms have been proposed, which do not necessarily exclude each other (Figure 1).

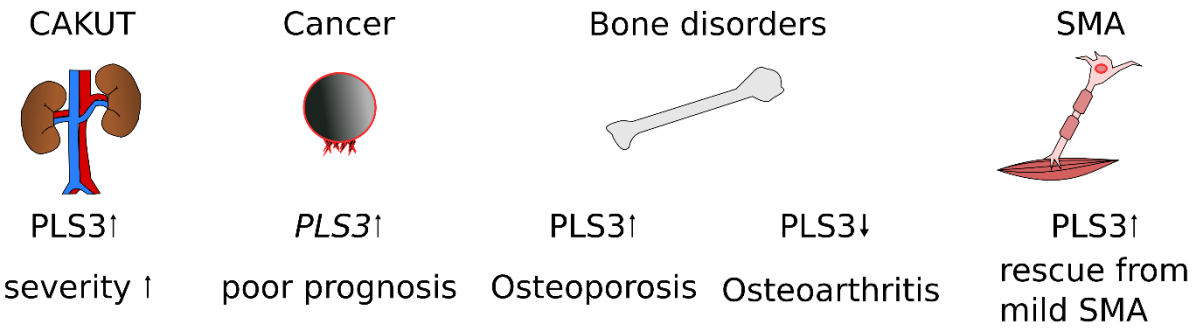


Figure 1: Alterations of the expression of *PLS3* influence the severity and prognosis in various diseases.

3.1.2. PLS3 and disorders of the musculoskeletal system

Genetic variants of *PLS3* cause osteoporosis with fractures in men and show the highest association with osteoporosis in women after menopause (van Dijk et al., 2013). In osteocytes, knock-out or downregulation of the gene is associated with early-onset osteoporosis (Makitie et al., 2020a; Makitie et al., 2020b; Makitie et al., 2020c; Tsolis et al., 2015). Increased levels of PLS3 in chondrocytes lead to a thickening of the cortical bone strength and are associated with osteoarthritis (Neugebauer et al., 2018; Tsolis et al., 2015). At least 27 genetic variants of *PLS3* with impact on osteoporosis have been identified; most of them are frameshift or nonsense mutations, which disrupt the function of the protein (Wolff et al., 2021).

3.1.3. PLS3 and neuromuscular disorders

PLS3 is associated with a broad spectrum of neuromuscular disorders, such as SMA, Amyotrophic Lateral Sclerosis [MIM: 105400] and *CHP1*-Associated Ataxia [MIM: 618438], where it acts as a genetic protective modifier (Janzen et al., 2019; Oprea et al., 2008; Walsh et al., 2020). SMA is characterised by progressive degeneration of α-motor neurons in the anterior ventral horns of the spinal cord. This ultimately results in symmetrical proximal muscle weakness and atrophy in the patients. The genetic cause of the SMA is deletion or mutation of the Survival of Motor Neuron 1 (*SMN1*, [MIM:

600354]) gene (Lefebvre et al., 1995). Interestingly, the genetic locus, which harbours *SMN1* in humans (5q11.2-q13.3) features a large inverted duplication with at least five duplicated genes (Figure 2) (Brzustowicz et al., 1990; Melki et al., 1990).

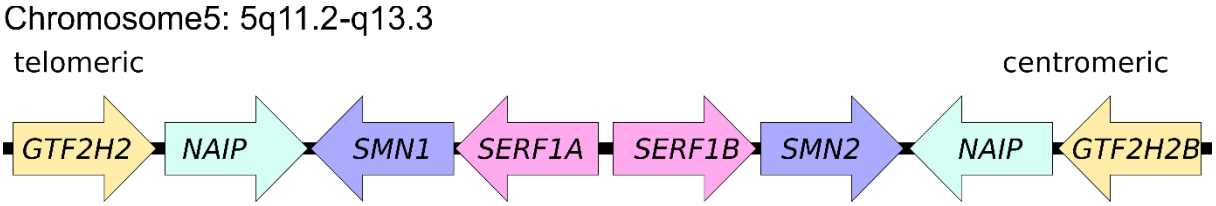


Figure 2: The inverted duplication at 5q112-q13q3. The 500 kb telomeric and centromeric region containing *SMN1* and its duplicated version *SMN2*, respectively (modified from Butchbach, 2021).

One of those genes is Survival of Motor Neuron 2 (*SMN2*, [MIM: 601627]). All *SMN1*-deleted SMA patients carry one or multiple copies of *SMN2*, which differs from *SMN1* in five silent nucleotide exchanges. A C to T transition at position +6 in exon 7 affects an exonic splicing enhancer, which acts as a *cis*-regulatory element (Burglen et al., 1996; Lefebvre et al., 1995; Lorson et al., 1999). This element is a hub for *trans*-acting proteins that are involved in the splicing machinery by recruiting serine-rich-like splicing factors (Cartegni and Krainer, 2002; Lorson et al., 1999). The nucleotide exchange leads to frequent skipping of exon 7 during the splicing process. By this, only 10% full-length SMN protein is transcribed from each *SMN2* copy compared to *SMN1* (Helmken et al., 2003). The remaining 90% mRNA lack exon 7 (Δ7-*SMN2*) and are translated into an unstable protein, which is rapidly degraded (Burnett et al., 2009; Lorson et al., 1999). Consequently, the severity of SMA, which varies from severe type I to adult type IV, inversely correlates with the copy number of *SMN2* (Feldkotter et al., 2002). The most severe type I (SMA1 [MIM: 253300]) has an onset in the first six months of life. Approximately 50% of patients suffer from this type of SMA. They are unable to sit or walk, their life expectancy is below two years, while nutritional and respiratory support is necessary. Most SMA1 patients carry two *SMN2* copies. The intermediate type II SMA (SMA2 [MIM: 253550]) manifests after six months of life. Patients are able to sit, but never achieve the ability to walk. They usually carry three *SMN2* copies. Milder affected type III (SMA3 [MIM: 253400]) patients are usually able to sit and walk, but often become wheelchair-bound with the progression of the disorder, starting after 18 months of life (Mercuri et al., 2018). Three to four SMN2 copies are usually found in SMA3 patients. The adult-onset form of SMA (SMA4 [MIM: 271150]) is characterised by a very late onset in the third decade of life and affected individuals show only very mild motoric dysfunction. Four to six *SMN2* copies are usually found in type IV patients

(Feldkotter et al., 2002; Wirth et al., 2006). Complete loss of the SMN protein is embryonically lethal (Schrank et al., 1997). However, it is poorly understood why motor neurons are selectively vulnerable to loss of SMN. Although the name “Survival of Motor Neuron” implies that the proteins primary function is maintenance of neurons, the protein is involved in a plethora of different cellular functions, such as ribonucleoprotein assembly, splicing, miRNA-biogenesis, trafficking into axonal compartments, local axonal translation, interaction with the cytoskeleton, endocytosis, autophagy, mitochondrial homeostasis, bioenergetics pathways, and the ubiquitin pathway, amongst others (Chaytow et al., 2018). This becomes especially clear, as the SMN gene product is ubiquitously expressed and in severe cases the dysfunction of SMN leads to multi-organ impairment (Hamilton and Gillingwater, 2013).

The function of SMN was originally described in the process of small nuclear ribonucleoprotein (snRNP) assembly (Burghes and Beattie, 2009; Pellizzoni, 2007). The snRNP is part of the spliceosome that catalyses the removal of introns from pre-mRNA. The SMN protein forms the macromolecular SMN complex together with several other proteins. This complex is involved in the biogenesis of snRNPs in the cytoplasm and trafficking into the nucleus. The assembly of snRNPs is more reduced in SMA mouse neural tissues, which might explain the severe effect of SMN dysfunction on motor neurons (Gabanella et al., 2005). Another explanation is given by a snRNP-independent motor neuron specific role of SMN. There seems to be a progressive shift of SMN localisation from the cytoplasm to the axon during development (Giavazzi et al., 2006). Several groups suggested that SMN interacts with mRNA-binding proteins (mRNPs) to maintain the localisation of axonal mRNAs into axons (Akten et al., 2011; Fallini et al., 2011, 2014; Hubers et al., 2011; Sanchez et al., 2013). Furthermore, motor neurons isolated from a severe SMA mouse model have a reduced average axon growth and the β-actin mRNA and protein levels are reduced in distal axons and growth cones (Rossoll et al., 2003). SMN is not only involved in the transport of mRNAs into axonal compartments, there is also evidence, that the protein has a role in local translation of mRNA transcripts through interaction with polyribosomes (Bernabo et al., 2017). Furthermore, SMN seems to be involved in the process of endocytosis (Hosseinibarkooie et al., 2016). Depletion of SMN in mice and the nematode *Caenorhabditis elegans* leads to impaired endocytosis and reduces the number of presynaptic docked vesicles. This led to disruption of the neuromuscular

junctions (NMJs), reduced synaptic transmission and synaptic vesicle recycling (Dimitriadi et al., 2016; Hosseinibarkooie et al., 2016).

3.1.4. PLS3 as genetic modifier of SMA

Severely affected SMA models are characterised by reduced axonal growth and connectivity at the NMJs. Neurotransmission and synaptic vesicle recycling are reduced (Ackermann et al., 2013; Kong et al., 2009; Rossoll et al., 2003). Loss of SMN is associated with impaired F-actin dynamics, most likely caused by delayed axonal transport of β-actin mRNA (Rossoll et al., 2003). As F-actin is also essential for endocytosis, overexpression of PLS3 counteracts major hallmarks of SMA pathology (Wu et al., 2016). The gene has been identified as a protective modifier in seven SMA-discordant families. All nine asymptomatic individuals in these families carry the causative *SMN1* deletion and the same *SMN2* copy number as their affected relatives and show increased levels of PLS3 (Heesen et al., 2016; Oprea et al., 2008). Overexpression of PLS3 restores endocytosis in SMA mice to a level, which is comparable to controls (Hosseinibarkooie et al., 2016). Furthermore, PLS3 is able to rescue axonal growth defects and motor neuron function in humans and various animal models (Hosseinibarkooie et al., 2016; Oprea et al., 2008).

In healthy individuals, expression of PLS3 was thought to be limited to solid tissues (Chafel et al., 1995; Lin et al., 1999). However, further investigations have shown that the gene is expressed in blood of about 5% of healthy individuals (Oprea et al., 2008). Asymptomatic *SMN1*-deleted individuals show an up to 40-fold upregulation of PLS3 in blood and lymphoblastoid cell lines derived from white blood cells compared to their affected siblings (Heesen et al., 2016; Oprea et al., 2008). Strikingly, in contrast to blood, in fibroblasts (FBs) from the same discordant siblings, no differences in PLS3 expression have been found, while induced pluripotent stem cell (iPSC)-derived spinal motor neurons (MNs) differentiated from these fibroblasts showed an upregulation of *PLS3* on RNA and protein level only in the asymptomatic but not symptomatic individuals (Figure 3) (Heesen et al., 2016).

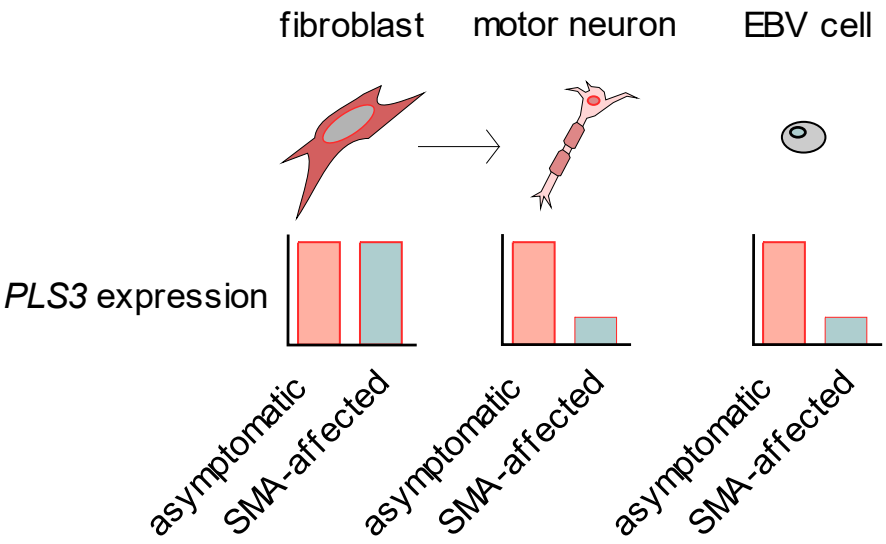


Figure 3: Expression of *PLS3* in various cell types. *PLS3* levels in FBs are increased without differences between asymptomatic (red) and SMA-affected individuals (green). iPSC-derived spinal MNs differentiated from the same FBs show differential expression of *PLS3*, which is upregulated in asymptomatic individuals. In Epstein-Barr Virus transformed lymphoblastoids (EBV cells) the gene is differentially expressed similar to the MNs.

3.1.5. PLS3 and cancer

The most prevalent malignancies of solid tissues are colorectal, prostate, breast, gastric and lung cancer. Overexpression of *PLS3* in circulating tumour cells (CTCs) is a common trait of all these cancer types and the protein was suggested as possible blood biomarker (Kujawski et al., 2015; Kurashige et al., 2019; Kuriyama et al., 2021; Sugimachi et al., 2014; Ueo et al., 2015; Xin et al., 2020; Yokobori et al., 2013). The most common cause of death in these malignancies are metastases in distant organs. Cell migration and tumour invasion are linked to changes in the actin cytoskeleton in a process called epithelial-mesenchymal transition (EMT). During this process, the cancer cells loose epithelial features and gain a mesenchymal phenotype (Thiery, 2002). EMT is an important factor for metastatic dissemination and acquisition of therapeutic resistance (Baum et al., 2008; Hugo et al., 2007; Polyak and Weinberg, 2009; Thiery and Sleeman, 2006; Yang and Weinberg, 2008). Elevated *PLS3* levels in CTCs contribute to a poor prognosis; however, it is still unclear, if the *PLS3* expression in CTCs is a trigger for EMT or a secondary effect of the cancer progression. For example, gastric cancer has the ability to migrate via the gastric wall into the adjacent tissue, which facilitates metastases to lymph nodes and distant organs after EMT (Kurashige et al., 2019). High *PLS3* expression levels in gastric cancer are associated with advanced tumour stages, cancer differentiation, tumour invasion depth, and

distant metastases. In addition to that, patients with increased expression of *PLS3* had a poor survival rate compared to low expressers, while suppression of *PLS3* lead to a decreased invasiveness of the cancer cells (Kurashige et al., 2019). In colorectal cancer, it has been found that the overexpression of *PLS3* is caused by a gain of the copy number of the region Xq23, the very same region that encodes *PLS3* (Sugimachi et al., 2014). Another study in colorectal cancer linked overexpression of *PLS3* to a genetic variant. The intronic *PLS3* polymorphism rs871773 C>T (X-114880523-C-T (GRCh37)) is associated with a low recurrence time of colorectal cancer in a sex-specific manner (Szkandera et al., 2013).

In addition to malignancies of solid tissues, overexpression of *PLS3* has been reported from malignancies of the hematopoietic and lymphatic system, such as acute myeloid leukaemia. Here, knock-down of *PLS3* increases the survival rate (Velthaus et al., 2019). Another malignancy of the hematopoietic system is the Sézary Syndrome (SS), a rare and aggressive variant of cutaneous T-cell lymphoma with a poor survival rate. This disease is characterised by the release of atypical CD4+ T-cells into the blood stream and the lymph nodes (see review by Kohnken et al., 2016). Those SS CD4+ T-cells show increased expression of *PLS3*, Twist Family BHLH Transcription Factor 1 (*TWIST1*) and GATA Binding Protein 6 (*GATA6*) compared to normal $CD4^+$ T-cells. *PLS3* has been suggested as a biomarker for SS CD4+ T cells associated with an unfavourable disease outcome. The upregulation was explained by hypomethylation of the promoter regions of all three genes, which are located on three different chromosomes. This points towards an epigenetic regulation (van Doorn et al., 2004; Wang et al., 2011; Wong et al., 2015). Hypomethylation is usually associated with gene activity, while silent genes are hypermethylated. In cancer tissue, a strong increase in hypomethylation is often found (Bedford and van Helden, 1987; Brothman et al., 2005; de Capoa et al., 2003; Ehrlich, 2002; Gama-Sosa et al., 1983).

The expression of *TWIST1* and *GATA6* are known to be driven by the important epigenetic transcriptional regulator nucleosome remodelling complex (NuRD) (Mohd-Sarip et al., 2017; Tremblay et al., 2018). Initially identified as transcriptional silencer, the NuRD complex is able to inhibit or activate gene transcription. The unique characteristic of NuRD is that it combines chromatin remodelling and protein deacetylase activity (Xue et al., 1998). The protein Chromodomain Helicase DNA Binding Protein 4 (CHD4, [MIM: 603277]), is a core member of the NuRD complex.

Interestingly, ChIPseq in therapy-resistant tumour initiating glioblastoma cells revealed a direct binding of CHD4 to the *PLS3* promoter (Chudnovsky et al., 2014).

3.1.6. PLS3, X-chromosomal inactivation and escape

Mary Lyon proposed in 1961 that in females one of the two X-chromosomes becomes inactivated to ensure gene dosage compensation between females (XX) and males (XY) (Lyon, 1961). In each of the female somatic cells one of the two X-chromosomes becomes randomly inactivated and forms the Barr body. Therefore, female somatic tissues are a mosaic of cells with inactivated X-chromosomes (X_i) of paternal or maternal origin (Migeon, 2007). All fully asymptomatic individuals from seven SMA-discordant families overexpress PLS3 and are female. This indicates a sex-specific mechanism of upregulation, such as escape from X-chromosomal inactivation (XCI). *PLS3* is indeed X-linked and a known facultative escape gene (Balaton et al., 2015; Carrel and Willard, 2005; Cotton et al., 2013; Yang et al., 2010).

The inactivation of the future X_i, is driven by several long non-coding RNAs (lncRNAs) that are encoded in a region called X-inactivation centre, located on the long arm of the human X-chromosome (Heard and Avner, 1994). One of those lncRNAs, which is essential for the XCI, is called X-inactive Specific Transcript (*XIST*, [MIM: 314670]). At the beginning of XCI, *XIST* is transcribed specifically from the future X_i. From there, the lncRNA spreads in *cis* along the X-chromosome and induces silencing by recruitment of histone-modifying enzymes (Cotton et al., 2014). Regions that are prone to be silenced become depleted of RNA Polymerase II (Chaumeil et al., 2006; Okamoto et al., 2004). The histone marks that can be found in inactivated regions are typical for heterochromatin. Additionally, inactive regions of the X_i are enriched with the repressive histone variant macroH2A (Changolkar et al., 2010). Finally, DNA methylation patterns correlate with the inactivation status of X-linked genes. Approximately 15% of X-linked genes escape XCI and are biallelically expressed. Another 15% of genes, including *PLS3*, variably escape in a tissue-specific manner (Balaton et al., 2015; Carrel and Willard, 2005; Cotton et al., 2013; Yang et al., 2010). Those escape genes show hypermethylation within the gene bodies, while promoter regions are hypomethylated (Berletch et al., 2015; Goto and Kimura, 2009; Kucera et al., 2011; Lister et al., 2013; Murakami et al., 2009; Schultz et al., 2015; Sharp et al., 2011). The mechanisms that trigger the escape of *PLS3* from XCI in a facultative manner are unknown.

3.1.7. The macrosatellite DXZ4 and escape from XCI

In recent years it became clear that our genome is partitioned into megabase scaled highly self-interacting regions called topologically associating domains (TADs) (Dixon et al., 2012). TADs are separated from each other by TAD boundaries, which suppress interactions between different TADs and are highly conserved among species, cell types and tissues (Dixon et al., 2012; Nora et al., 2012). TAD boundaries are often enriched for binding sites for the CCCTC-Binding Factor (CTCF) (Dixon et al., 2012). Interestingly, the X_i shows a chromosomal conformation, which is distinctive from the active X-chromosome (X_a), as it is partitioned into two massive megadomains (0-115 Mb and 115-155.3 Mb). The X-chromosome harbours a unique macrosatellite, called *DXZ4*, which is located next to the boundary of the two megadomains (ChrX: 114,867,433 – 114,919,088 bp) (Rao et al., 2014). *DXZ4* has a highly variable repeat number, which was estimated as 50-100 copies of a 3 kb repeat monomer (Giacalone et al., 1992). This macrosatellite is essential for XCI, as its deletion leads to a de-condensation of the X_i in mice and a disruption of the two megadomains (Bonora et al., 2018; Darrow et al., 2016; Figueroa et al., 2015). Surprisingly, in males and on the female X_a, *DXZ4* features heterochromatic epigenetic features. It is hypermethylated and does not bind CTCF. On the X_i, the macrosatellite is euchromatic, hypomethylated and binds CTCF (Rao et al., 2014). It has been reported, that the mouse homologue of the macrosatellite (*Dxz4*) has a much smaller copy number of only seven repeats (Horakova et al., 2012). Furthermore, the escape of *PLS3* is highly variable in human but not in mouse females (Carrel and Willard, 2005; Yang et al., 2010). In both species *PLS3/Pls3* is located next to *DXZ4/Dxz4* (Horakova et al., 2012). It was suggested that the highly variable copy number and unusual chromosomal conformation of *DXZ4* influence the expression of neighbouring genes, such as *PLS3* (Horakova et al., 2012).

3.2. Part II: The impact of PLS3 on biomarkers of SMA

3.2.1. PLS3 as biomarker

A biomarker must show the ability to discriminate affected individuals from those that are not or will not be affected by the disease of interest and must strengthen the correlation agreement between predicted and observed frequency of the outcome according to a risk prediction model (Wang, 2011). Furthermore, a biomarker should be easily and safely accessible at low costs in order to be useful in medicine. In addition to that, the biomarker should give advances to already established methods; otherwise, it will not be used in day-to-day practice (Wang, 2011).

Usually, biomarkers are measured in body fluids, such as blood or urine as they are easily accessible. Several studies tried to establish PLS3 as a biomarker (Fedou et al., 2021; Tang et al., 2010). Some studies focus on the analysis of *PLS3* polymorphisms, as the mutational burden, the occurrence of specific single nucleotide variants (SNVs) compared to a reference can imply important information about the progression of a disease (Lyberopoulou et al., 2015; Makitie et al., 2020b; Makitie et al., 2020c; Ning et al., 2014). Relevant information can be extracted from normal tissue to predict the potential to develop a disease. Especially in malignancies, the mutational burden of cancer cells, such as CTCs is an important source of information. In CTCs from colorectal cancer patients, especially the SNV rs6643869 seems to be of much value as stage-specific molecular predictor of tumour recurrence (Lyberopoulou et al., 2015; Ning et al., 2014). A more recent study in childhood-onset osteoporosis patients analysed serum concentration of 192 microRNA (miRNAs) in patients with or without *PLS3* mutations. They identified a population of seven miRNAs with a specific expression pattern. None of these were previously associated with *PLS3* (Makitie et al., 2020b; Makitie et al., 2020c).

Other biomarker studies are more interested in the general expression levels of PLS3 in different disease conditions. Increased levels were found in foetal kidney cells and the abundance of the protein is positively associated with the severity of congenital anomalies of the kidney and urinary tract (CAKUT) (Fedou et al., 2021). However, kidney cells are not easily accessible. A very interesting and promising approach is to use PLS3 as biomarker to differentiate disorders with a similar phenotype. As described above, *PLS3* overexpression is linked to SS and was suggested as

biomarker (Wong et al., 2015). In addition to SS, there is another common form of cutaneous T-cell lymphomas called mycosis fungoides. The two malignancies share similar disease etiologies and the diagnosis can be challenging (Larocca and Kupper, 2019). In an interesting study, the expression of *PLS3* has been measured in peripheral blood mononuclear cells during disease progression to differentiate the two malignancies. They found a 400-fold difference in the expression between mycosis fungoides and SS patients and an increase of the expression difference during disease progression. More importantly, the expression levels of *PLS3* decreased during treatment with chemotherapy and bone marrow transplantation (Tang et al., 2010). In a follow up study, no *PLS3* mutations were identified that could explain the expression patterns; however, significant hypomethylation of *PLS3* was observed in the individuals with increased expression of *PLS3*. A loss of CD26 was associated with gain of *PLS3* (Jones et al., 2012). In a similar approach, expression of *PLS3* was used to distinguish pancreatic ductal adenocarcinoma from diffuse large B-cell lymphoma (Xiong et al., 2021).

3.2.2. Impact of PLS3 on biomarkers for SMA

Intensive effort over multiple decades led to the development of therapies for SMA using either small molecules that alter *SMN2* splicing in order to increase the amount of functional SMN, antisense oligonucleotides (ASOs) or adeno associated virus (AAV)-based gene therapy (Scoto et al., 2017; Wirth, 2021). One of these SMA treatment options, an ASO known as Spinraza (Nusinersen) has been FDA and EMA approved for SMA therapy (Scoto et al., 2017). It blocks an intronic silencer in *SMN2* transcripts and promotes exon 7 inclusion (Rigo et al., 2014). Nevertheless, SMA is a multi-organ disease and the effects of low levels of SMN in treated patients will be seen in the years to come. The SMN protein levels in most severely affected tissues are not easily measurable, as the main tissue, motor neurons, are not directly accessible. Additionally, motor score tests, such as the Modified Hammersmith Functional Motor Scale (MHFMS) can be inappropriate for children under 30 months and may not be comparable to motor score test that are specifically designed for young infants(Main et al., 2003). For these reasons, secure progression and treatment monitoring using blood biomarkers will be of large benefit for SMA patients that are often fragile and immobile. Several studies have shown that the disease and treatment progression can be reliably monitored by measurement of plasma protein levels of SMN and other

plasma proteins (Arnold et al., 2016; Arnold et al., 2014; Bonati et al., 2017; Crawford et al., 2012; Czech et al., 2015; Finkel et al., 2012; Iyer et al., 2017; Kolb et al., 2016; Viswambharan et al., 2017; Wadman et al., 2016).

One large study was the BforSMA project, which identified candidate proteins, transcripts and metabolite biomarkers in an unbiased approach. The study evaluated blood and urine protein analytes in SMA-affected children and age-matched controls. The analytes were associated with MHFMS scores, which is a commonly used test setup for SMA disease severity (Crawford et al., 2012; Finkel et al., 2012). They identified 200 proteins that correlated or anti-correlated with disease severity. The plasma levels of ten of those proteins were analysed in a severe Δ7-SMA mouse model (Arnold et al., 2016). In this mouse model, overexpression of PLS3 restores the functions of the motor neurons and the NMJ, however, the survival rate of the mice was not prolonged, due to multi-organ dysfunctions caused by low amount of endogenous SMN protein (Ackermann et al., 2013). Asymptomatic individuals are usually found in families with SMA-affected individuals that suffer from mild forms of SMA. To gain more insight how PLS3 is modifying the disease, we recently developed an intermediate Taiwanese SMA mouse model by injecting pre-symptomatically a suboptimal low-dose SMN-ASO (30 µg), which resembled the phenotype of SMA2 in patients. By this, the survival of the animals was prolonged from 16 days to four weeks (26 ± 9.48 days). Additional *PLS3* overexpression from a transgenic allele prolonged the survival from 26 to more than 250 days in more than 60% of mice (Hosseinibarkooie et al., 2016).

4. Aims of the study

4.1. Part I: Regulation of PLS3 in health and disease

Little is known about the regulation of *PLS3* and multiple different mechanisms that do not exclude each other have been proposed. The main aim in the first part of the study was to identify mechanisms that contribute to the regulation of *PLS3* in various cell types.

1. We aimed to analyse transcriptome data from fibroblasts, lymphoblastoid cells and iPSC-induced spinal motor neurons from asymptomatic and SMA-affected individuals as well as healthy controls. Our first objective was to reproduce and validate previously described *PLS3* expression patterns.

All fully asymptomatic individuals from seven SMA-discordant families are female and *PLS3* is a known facultative escape gene (Balaton et al., 2015; Carrel and Willard, 2005; Cotton et al., 2013; Yang et al., 2010).

2. We aimed to investigate, if differences in the expression of *PLS3* between male and female siblings can be explained by an escape of *PLS3* from XCI.

Furthermore, *PLS3* is located next to the macrosatellite *DXZ4*, which has an unusual chromosomal conformation and a highly variable copy number (Giacalone et al., 1992; Rao et al., 2014). We hypothesise that the copy number of *DXZ4* influences the expression of *PLS3*.

3. We aimed to measure the copy number of *DXZ4* by Molecular Combing and to compare the results to the expression levels of *PLS3* in lymphoblastoid cells.

However, the gene is upregulated in blood of 5% of the healthy population in both genders, and there are rare cases of increased *PLS3* expression in males. We assume that there are at least two different mechanisms that contribute to the upregulation of *PLS3* (Oprea et al., 2008).

4. Our last aim was to identify transcription factors or transcriptional regulators that influence the expression of *PLS3* by analysis of transcriptome data and further analysis in cell culture experiments.

4.2. Part II: The impact of PLS3 on biomarkers of SMA

The pilot study of the BforSMA project identified candidate plasma protein biomarkers for SMA in an unbiased approach. Ten of those putative biomarkers were analysed in the context of a severe Δ7-SMA mouse model (Arnold et al., 2016). We have previously developed an intermediate mouse model, which resembles the situation of a mild SMA pathology (SMA3) in humans. Our model was characterised by an increased survival rate from 16 days to four weeks (26 ± 9.48 days), while overexpression of *PLS3* from a transgenic allele prolonged the survival from 26 to more than 250 days in more than 60% of mice (Hosseinibarkooie et al., 2016).

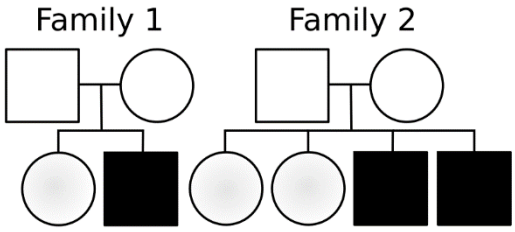
1. We aimed to analyse and evaluate the plasma concentrations of SMN, and six biomarkers from the BforSMA biomarker panel, namely COMP, DPP4, SPP1, CLEC3B, VTN, and AHSG in our intermediate SMA mouse model.
2. Furthermore, our objective was to assess whether *PLS3* overexpression or low-dose systemically administered *SMN*-ASO-therapy have any impact on these biomarkers.

5. Results

5.1. Part I: Expression of PLS3 in health and disease

5.1.1. PLS3 escapes from XCI in iPSC-derived spinal motor neurons

To validate previous results about the expression of *PLS3* in different tissues, we analysed transcriptomes from FBs of two SMA-discordant families (Figure 4, Table 1). We compared the results to transcriptomes from iPSC-derived spinal MNs that were previously reprogrammed from the same fibroblasts.



EBV cell line	AS f3	SMA m2	AS f1	AS f2	SMA m10	SMA m8
ID (Fibroblasts)	FB1	FB2	FB3	FB4	FB5	FB6
ID (Motor neurons)	MN1	MN2	MN3	MN4	MN5	MN6
DXZ4 copy no.	91/49	91	110/60	110/44	60	44
rs2108099	C/T	T	C/T	C/T	T	T
rs871774	C/T	T	C/T	C/T	T	T
Average SNV ratio	0.4		0.47	0.49		

Figure 4: Pedigrees of two SMA-discordant families. Both families comprise affected (black) and asymptomatic (grey) individuals, all carrying homozygous *SMN1* deletions and four *SMN2* copies.

Next, we performed a principal component analysis. The spinal MNs cluster in two non-overlapping groups (Figure 5A). In the fibroblasts, the asymptomatic individuals (red data points) cluster together, while the SMA-affected individuals (green data points) are more variable, but distinguishable from the asymptomatic individuals (Figure 5B).

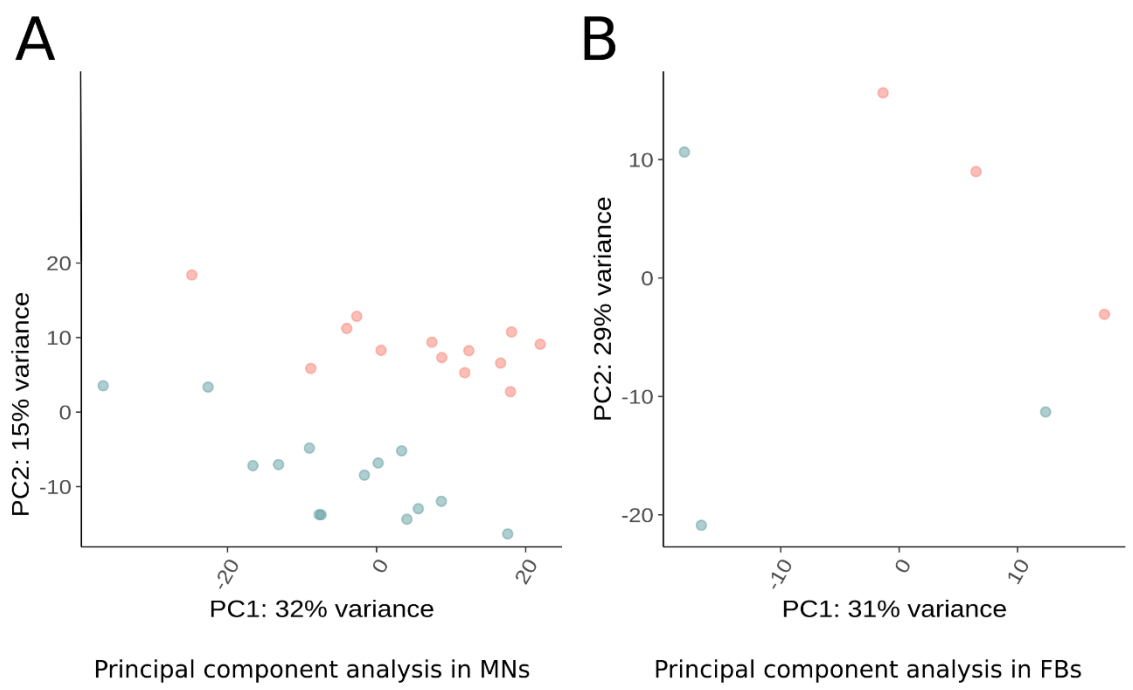


Figure 5: Principal component analysis. (A) Transcriptome data of MNs of asymptomatic (red) and SMA-affected individuals (green) cluster in two distinguishable groups. (B) Transcriptome data of FBs of asymptomatic (red) and SMA-affected individuals (green) cluster in two distinguishable groups.

The transcriptome data of FBs, EBV cells and spinal MNs were performed using KALLISTO and DESeq2 (Bray et al., 2016; Love et al., 2014). This pipeline uses 'Reads per Kilo Base per Million Mapped Reads' (RPKM) as method for transcriptome isoform quantification and normalisation. RPKM for each given gene can be calculated by equation 1 (see methods section 7.2.2.3) (Bray et al., 2016; Conesa et al., 2016; Love et al., 2014).

The analysis of FB transcriptomes of the six siblings showed low expression levels of *SMN*. This was expected, since all siblings share the homozygous deletion of *SMN1* as well as the same copy number of *SMN2* (Figure 6A, B). Furthermore, all siblings show strong expression of *PLS3* without differences between the SMA-affected and asymptomatic individuals confirming previously described expression patterns (Figure 6C, D) (Heesen et al., 2016; Oprea et al., 2008). In general, 17 genes were differentially expressed. The called genes include pseudogenes, such as *USP32P1*, however, no X-linked genes.

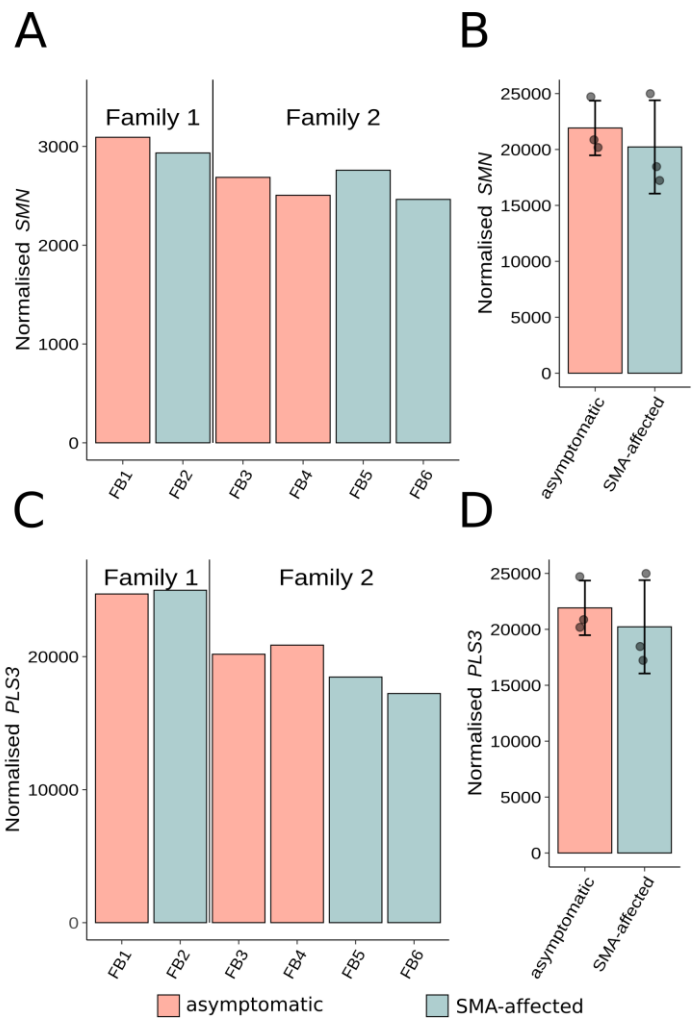


Figure 6: Transcriptome analysis of fibroblasts. Gene expression levels of *SMN* (A, B) and the genetic modifier *PLS3* (C, D) in FB transcriptomes from all six siblings of the two SMA-discordant families measured as RPKM for whole identified transcripts. Asymptomatic females are shown in red, while SMA-affected male siblings are shown in green. No differences between asymptomatic and SMA-affected individuals have been found. Data are represented as mean ± standard deviation (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$).

Next, we analysed the transcriptomes obtained from spinal MNs, which were differentiated from the same FB cell lines of the two SMA-discordant families (Heesen et al., 2016). For family 2, we obtained transcriptomes from two clones per individual sequenced in triplicates. For family 1, RNA from each of the two siblings was sequenced in triplicates. As expected, there were no differences in the expression of *SMN* (Figure 7A, B). However, we confirmed differential expression of *PLS3* in both families and found an approximately 1.12-log2 fold change, which means a doubling of the *PLS3* amount in the asymptomatic individuals in both families (Figure 7C, D).

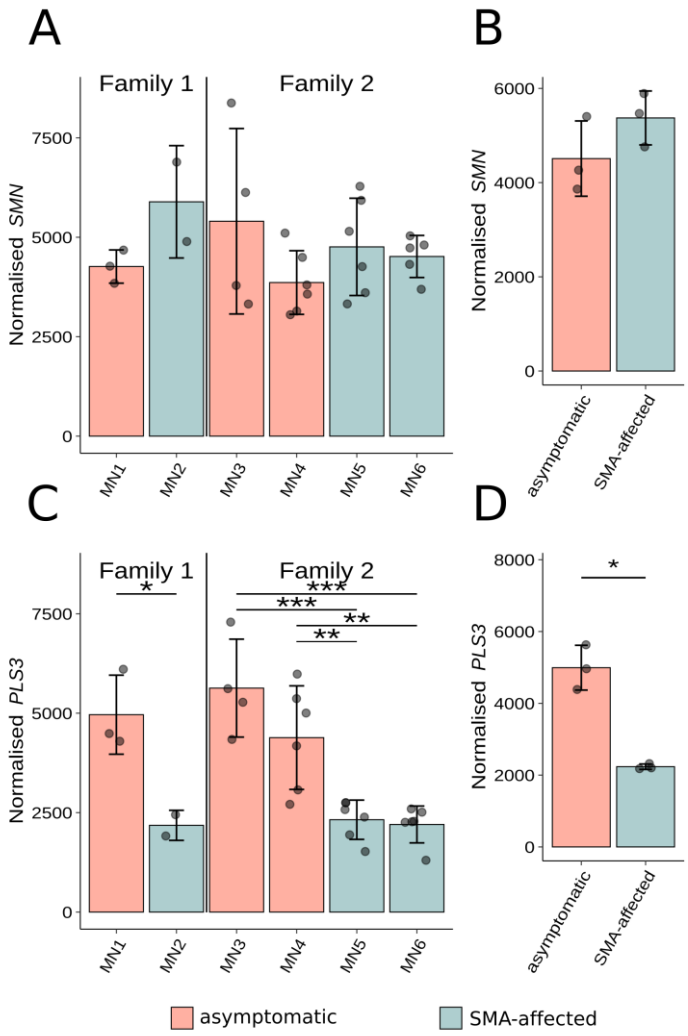


Figure 7: Transcriptome analysis of spinal motor neurons. Gene expression levels of *SMN* (A, B) and the genetic modifier *PLS3* (C, D) in transcriptome data of spinal MNs from all six siblings of the two SMA-discordant families measured as RPKM for whole identified transcripts. Asymptomatic females are shown in red, while SMA-affected male siblings are shown in green. No differences in the expression of *SMN* between asymptomatic and SMA-affected individuals have been found. We validated differential expression of *PLS3*. Data are represented as mean ± standard deviation (**P* < 0.05; ***P* < 0.01; ****P* < 0.001).

A comparison of all samples from SMA-affected siblings to the asymptomatic siblings identified 80 significant differentially expressed genes, including 61 X-linked genes (Figure 8, Table 2).

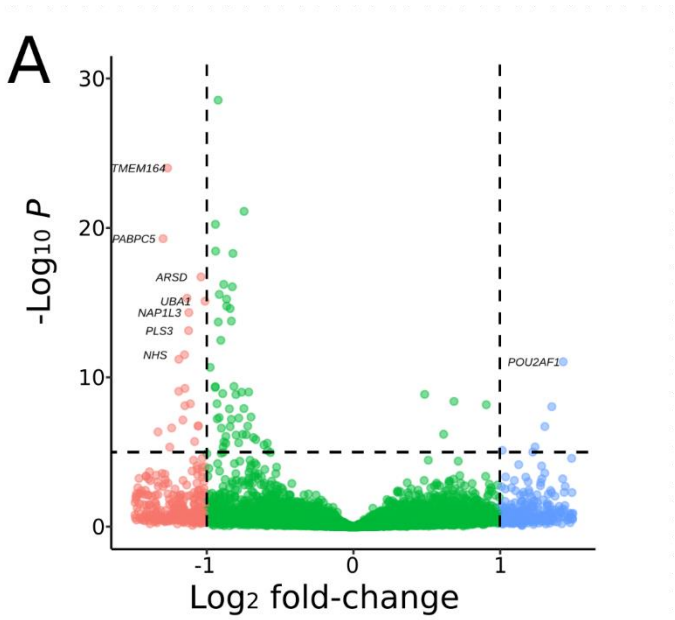


Figure 8: Differential expression analysis of transcriptomes of spinal motor neurons. Asymptomatic and SMA-affected individuals from both SMA-discordant families were compared. The volcano plot shows 80 differentially expressed genes, including *PLS3*, identified in transcriptomes of iPSC-derived spinal MNs. Genes with a $\log_2$ fold-change ≤ -1 are overexpressed in asymptomatic individuals and are given in red. Genes that are overexpressed in SMA-affected individuals are given in blue.

With our analysis of FB and MN transcriptomes, we were able to confirm previously described expression patterns of *PLS3* (Oprea et al., 2008). Next, we investigated if the difference in the expression levels of *PLS3* would be explained by an escape from XCI, as *PLS3* is a known facultative escape gene (Balaton et al., 2015; Carrel and Willard, 2005; Cotton et al., 2013; Yang et al., 2010). The iPSC-derived spinal MNs were differentiated from single-cell clones. As previously reported, the X-inactivation (XI) status is inherited from each clone to the daughter cells, which means that in each cell of the population the same X-chromosome of paternal or maternal origin is inactivated (Geens and Chuva De Sousa Lopes, 2017). We took advantage of this and identified escape genes from the transcriptomes by calling X-chromosomal heterozygous biallelic SNVs. Each biallelic SNV indicated that the respective gene escaped XCI, however, the shot-gun approach of short-fragment sequencing is error prone. Therefore, we filtered for a read depth larger than 100 reads. Finally, we selected only SNVs, which showed a SNV ratio between 0.1 and 0.9, which is a commonly used cut-off for the definition of escape genes in literature (Balaton and Brown, 2016; Calabrese et al., 2012; Carrel and Willard, 2005; Cotton et al., 2013, 2014; Yang et al., 2010; Zhang et al., 2013). This means that only genes, which were transcribed in a ratio between 1 by 10 or higher, were considered as true escape genes. In addition to that, we discarded all SNVs that belong to genes from the

pseudoauttosomal regions (PAR1 ChrX: 1-2,699,520, PAR2 ChrX: 154,931,044-160,000,000). By this method, we identified 147 SNVs referring to 107 biallelically expressed escape genes, including *PLS3* in both families (Figure 9A, Table 3).

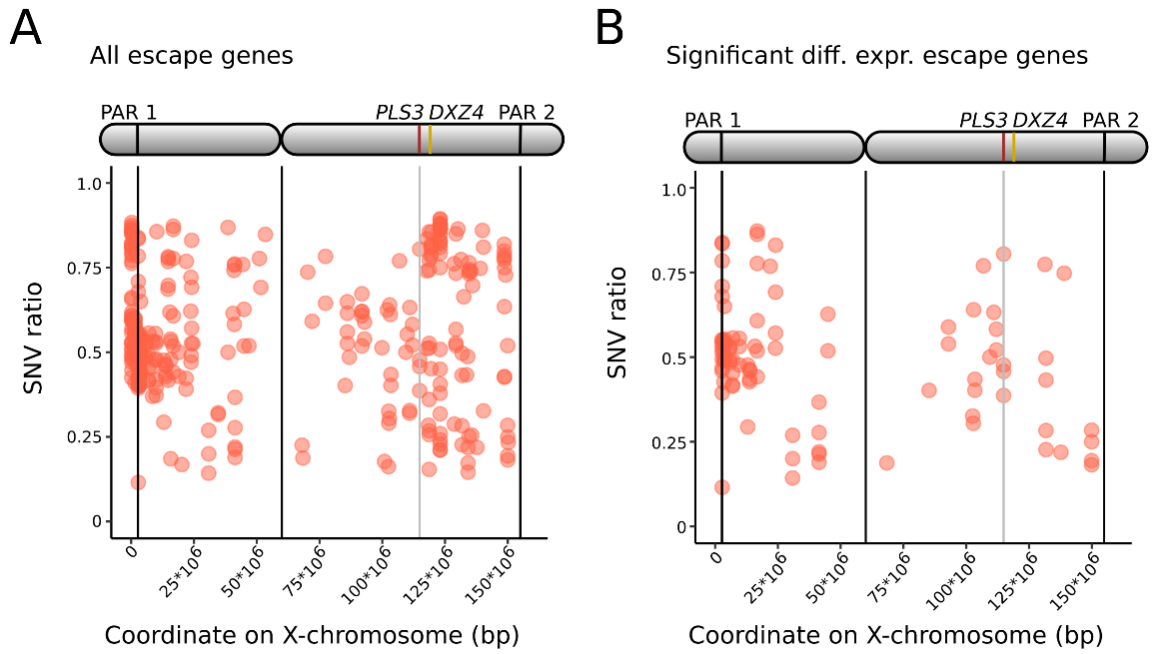


Figure 9: Escape genes in spinal motor neurons. (A) Escape genes (including *PLS3*) were identified in transcriptomes of spinal MNs by filtering exonic biallelic SNVs with an SNV ratio of 0.1 to 0.9. The escape genes were distributed along the X-chromosome. (B) Location and SNV ratio of escape genes that were significant differentially expressed, including *PLS3*.

The human X-chromosome contains of about 1,200 coding genes. In our analysis, we found that about 10% of X-linked genes escaped the XCI. Since we were only able to identify genes that contained exonic biallelic SNVs with our method, we concluded that this number seemed plausible. Not all biallelically expressed genes are necessarily differentially expressed, as the expression of the genes may be regulated by other mechanisms. Therefore, we compared our list of biallelically expressed genes with our list of significant differentially expressed genes. We found 35 differentially expressed escape genes, including *PLS3* (Figure 9B, Table 3).

The SNVs were distributed along the X-chromosome. In the male siblings, biallelic SNVs belonging to 15 genes were found outside of the pseudoautosomal regions (Figure 10A, B, C, Table 4). The majority of them is either a pseudogene or multicopy gene and can therefore be discarded. Interestingly, the same genes were called in males from both families, indicating that indeed those SNVs do not belong to escape genes.

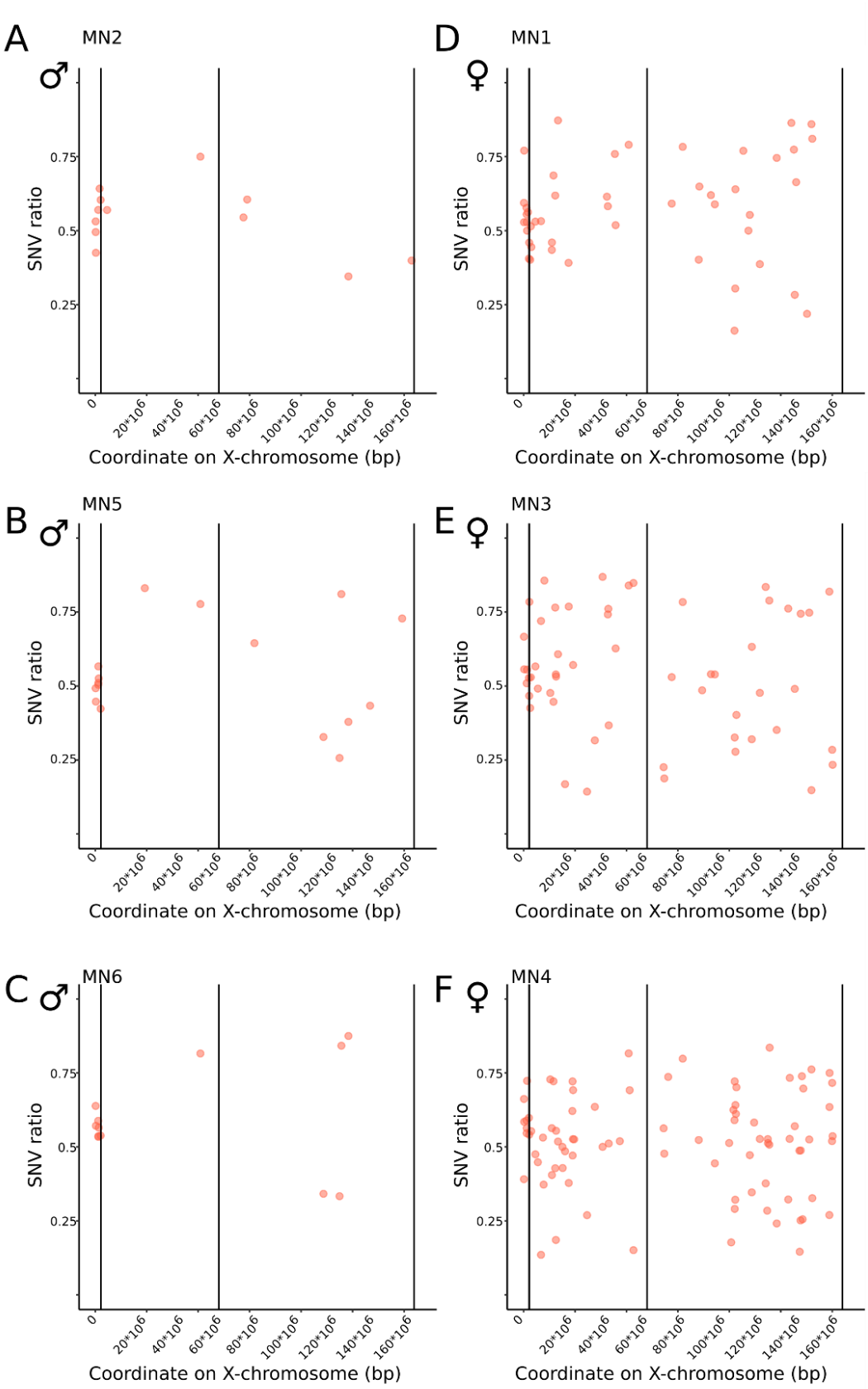


Figure 10: Biallelic, exonic SNVs identified in transcriptomes from spinal MNs. (A, B, C) The SNVs are distributed along the X-chromosome in the three male samples. For each SNV, the SNV ratio is given. (D, E, F). In the three female samples (MN1, MN3, and MN4), 15 genes were identified outside of the pseudoautosomal regions. The majority are pseudogenes or multicopy genes.

5.1.2. Comparison to other escape gene studies

We compared our list of escape genes with other studies. A study by Zhang et al. lists 114 escape genes identified in transcriptomes of B lymphocytes (Zhang et al., 2013). Carrel and Willard published a list of escape genes measured in 40 human FBs (Carrel and Willard, 2005). Cotton et al. analysed the methylation patterns of human leukocytes from 1,800 females and identified escape genes across 27 tissues (Cotton et al., 2015). Balaton et al. published a meta-study comparing previously published results (Balaton et al., 2015; Carrel and Willard, 2005; Cotton et al., 2015). That study lists the escape behaviour of 1,145 X-linked genes; 107 genes were escape genes (including *PLS3*). Katsir and Linial identified 75 escape genes in single-cell transcriptomes from FBs (Wainer Katsir and Linial, 2019). We compared our list of escape genes with the previously published gene lists (Figure 11). Interestingly, each list of escape genes contains approximately 20% of genes that are unique to each study. Our study identified 60 escape genes (20.1%) that are not found in the other lists. The meta-study by Balaton et al. includes 58 (19.4%) escape genes that were not found by the other groups. Katsir and Linial identified 41 (13.7%) escape genes, which were unique to their list and Zhang et al. listed 65 unique genes (21.7%). All escape gene lists together identified 299 different escape genes. Only seven genes (2.3%) were found by all escape gene studies.

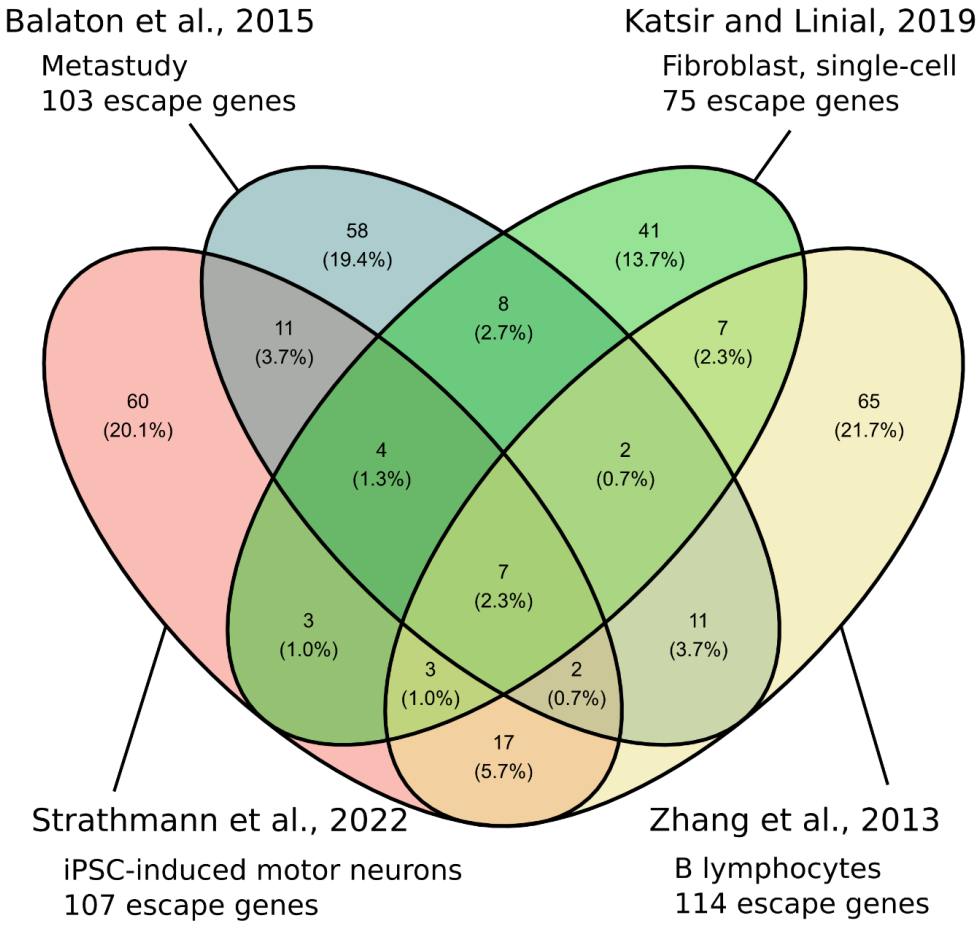


Figure 11: Comparison of escape gene lists of four different studies.

5.1.3. DXZ4 copy number correlates with PLS3 expression in females

PLS3 is located on the X-chromosome next to the unique macrosatellite *DXZ4*, which is characterised by a highly variable copy number estimated as 50 to 100 copies of a 3 kb monomer (Giacalone et al., 1992). On the X_i, *DXZ4* is characterised by euchromatin, while the X_a consists of heterochromatin (Chadwick, 2008). We hypothesised that the unusual chromosomal conformation and the highly variable copy number of *DXZ4* influences the expression of neighbouring genes, such as *PLS3*.

To investigate this, we performed a pilot study using 22 whole-genome sequencing data sets from asymptomatic and SMA-affected males and females with various *PLS3* expression as well as ten data sets from unrelated controls (Table 1). For each dataset, we calculated the average read depth of the genetic locus that contains *PLS3* and *DXZ4* (ChrX:114,827,819-114,885,179, GRCh37/ hg19). To normalise the data sets and to compare the resulting data, we calculated the average read depth of

chromosome 22, as this is the smallest chromosome. This average was used to normalise the read depths of *PLS3*/*DXZ4*. In addition to that, we took into account that males are hemizygous and thus have only one allele of *PLS3*. We plotted the normalised read depths of the locus that contains *PLS3* and *DXZ4* and found that *PLS3* showed a very constant read depth over the course of the coding region. However, the *DXZ4* locus (ChrX:114,959,000-115,006,000, GRCh37/ hg19) showed a region of variable read depth (Figure 12A). We assumed that this variable region consists of *DXZ4* repeat monomers that were aligned by the Burrows-Wheeler Alignment (BWA) algorithm to the reference genome, which was not able to align the fragments in the correct order, due to the shot-gun approach of DNA sequencing (Li and Durbin, 2009). We used the average read-depth of this variable region in each normalised sample to estimate the copy number of *DXZ4*. Next, we compared the estimated *DXZ4* copy numbers to *PLS3* expression levels previously measured in EBV cells by qRT-PCR and found a significant linear correlation ($P \approx 0.005$, $\rho \approx 0.51$) (Figure 12B) (Oprea et al., 2008).

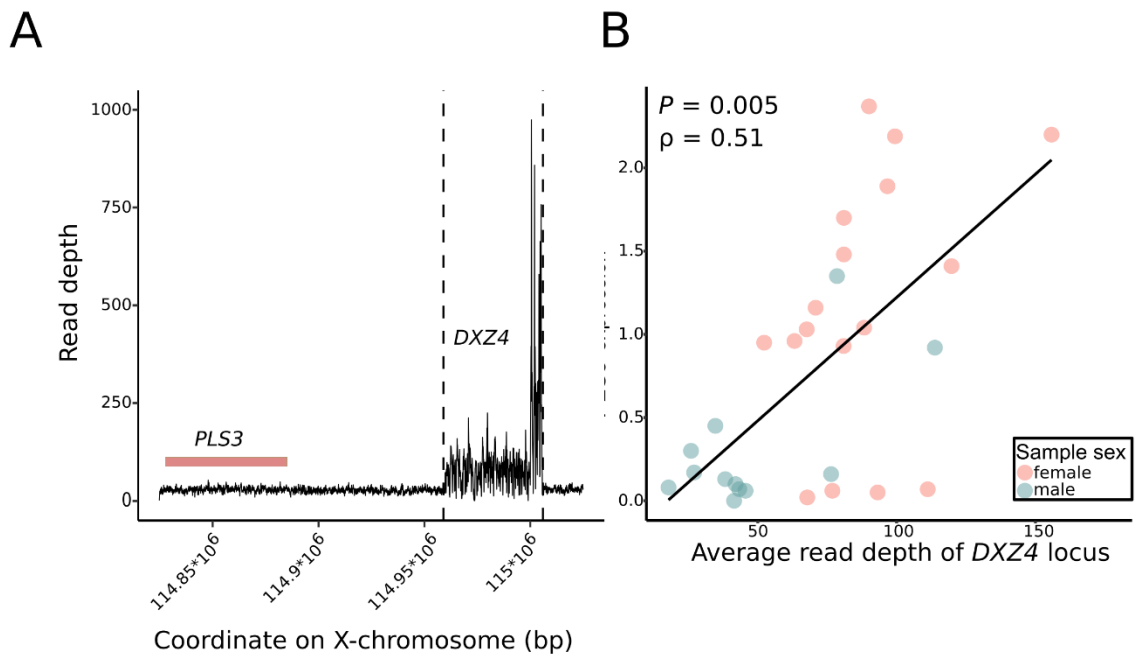


Figure 12: Estimation of the *DXZ4* copy number by bioinformatic analysis of the read depth in whole-genome sequencing data. (A) The read depth of 32 whole-genome sequencing data sets was measured and normalised. The *PLS3* locus shows a constant read depth, while the *DXZ4* locus shows a region that is highly variable. The average read depth of that region was used to estimate the copy number of *DXZ4*. (B) The average normalised read depth of *DXZ4* was measured in 32 whole-genome data sets and compared to expression levels of *PLS3* measured by qRT-PCR. A linear relationship was determined by Spearman's Correlation Coefficient test. (*P < 0.05; **P < 0.01; ***P < 0.001).

We have shown by this method, that there is indeed a linear relationship between the copy number of *DXZ4* and the expression levels of *PLS3*. However, we were not able

to discriminate the two *DXZ4* alleles in females, which might differ in their copy numbers. Furthermore, the copy numbers were rough estimates, while the real copy numbers were still unknown, which allowed no comparison to literature. To overcome these limitations and to measure the exact copy number of both *DXZ4* alleles, we applied a method called Molecular Combing (Michalet et al., 1997). We cultivated EBV cells of six asymptomatic and eleven SMA-affected females, as well as eight SMA-affected males, which have shown various *PLS3* expression levels in previous experiments. All siblings of both SMA-discordant families were included in this experiment (Table 1). We extracted RNA from each EBV cell line and measured the expression of *PLS3* by qRT-PCR. The inclusion of an internal reference allowed us to compare the expression levels of all samples against each other. Samples that showed *PLS3* levels of 1/10th or less compared to the highest value were defined as low expressers (Figure 13).

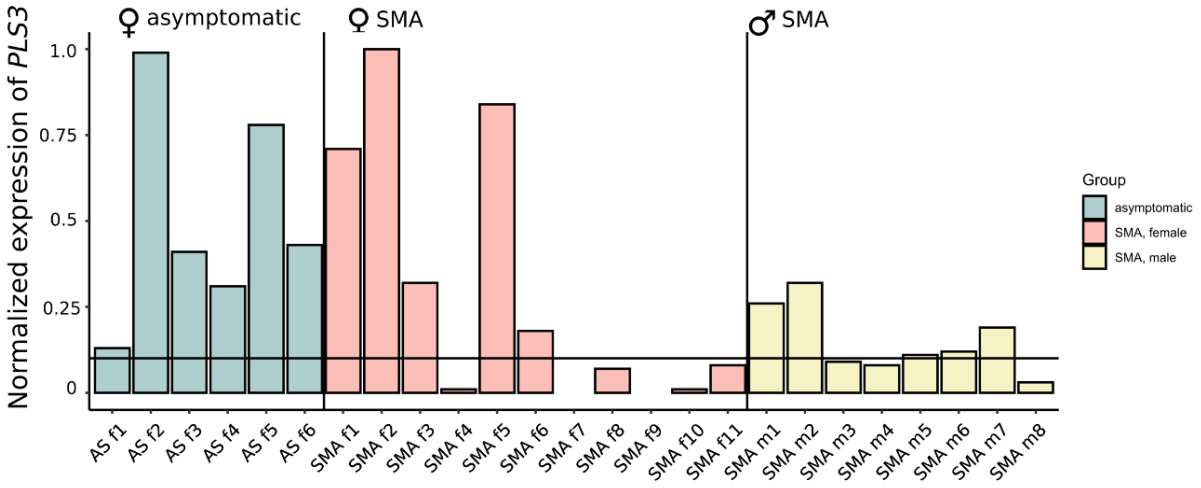


Figure 13: Expression of *PLS3* in asymptomatic and SMA-affected individuals. *PLS3* expression levels of 17 female and 8 male EBV cell lines measured by qRT-PCR. Samples are ordered by phenotype as female asymptomatic individuals (AS f, green), SMA-affected females (SMA f, red) and SMA-affected males (SMA m, yellow). Samples that express 10% or less compared to the strongest expresser were defined as *PLS3* low expressers for further analysis.

In parallel, we extracted high molecular weight genomic DNA from each sample. The genomic DNA was then solved, linearised on vinyl-silane coated coverslips and hybridised with custom made fiber probes. Each *DXZ4* repeat monomer of 3 kb was covered by two differently labelled fiber probes of 1.1 kb length. The regions upstream and downstream of *DXZ4* were covered by two additional fiber probes, allowing the selection of only intact DNA stretches covering the entire *DXZ4* region (Figure 14A).

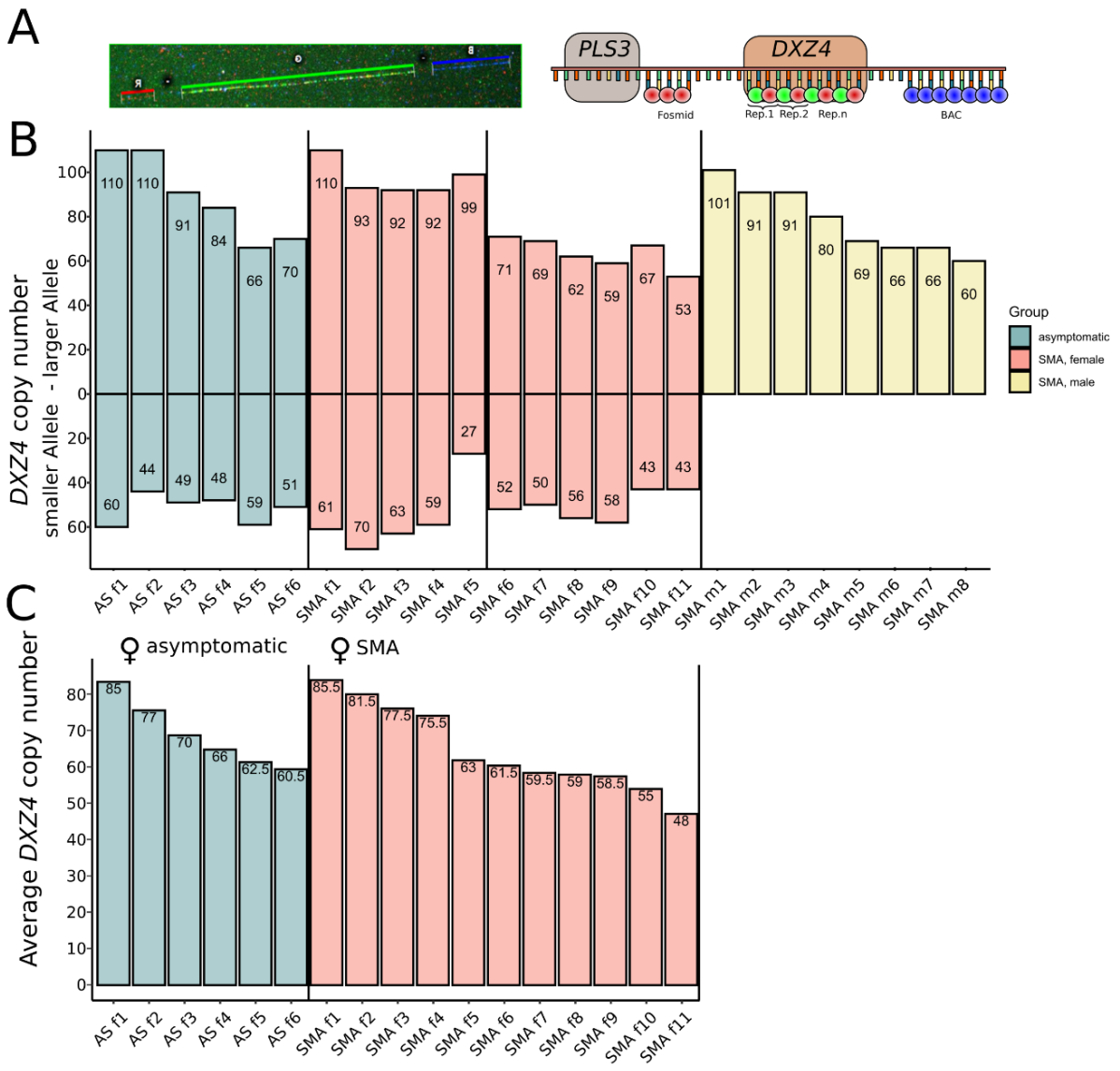


Figure 14: Measurement of *DXZ4* copy numbers by Molecular Combing. (A) The fiber probes used for Molecular Combing consist of two red and green 1.1 kb fiber probes covering each *DXZ4* repeat and two fiber probes covering the up- and downstream regions. (B) Copy numbers of *DXZ4* measured in EBV cells. The larger allele in females is shown in the upper half of the figure and the smaller allele in the lower half of the figure. (C) The average *DXZ4* copy number in females. Samples are ordered by phenotype as female asymptomatic individuals (AS f, green), SMA-affected females (SMA f, red) and SMA-affected males (SMA m, yellow).

The linearised and hybridised DNA fibres were scanned and the length of the *DXZ4* macrosatellites was determined manually using the FiberStudio Software. Only complete signals that consisted of all three regions were incorporated in the further analysis. The largest detected *DXZ4* copy numbers have about 110 repeats, while the smallest consisted of 27 repeat monomers (Figure 14B and figure 15). Interestingly, of the 17 female samples there was only one sample (SMA f9) with two *DXZ4* alleles of similar size. All other samples showed two clearly distinguishable *DXZ4* alleles. We calculated the average copy of *DXZ4* in the female samples and found that the largest

average copy numbers (85 repeats) were approximately twice as big as the smallest (48 repeats) (Figure 14C). Our findings underline the high variability of *DXZ4* copy numbers in the human population.

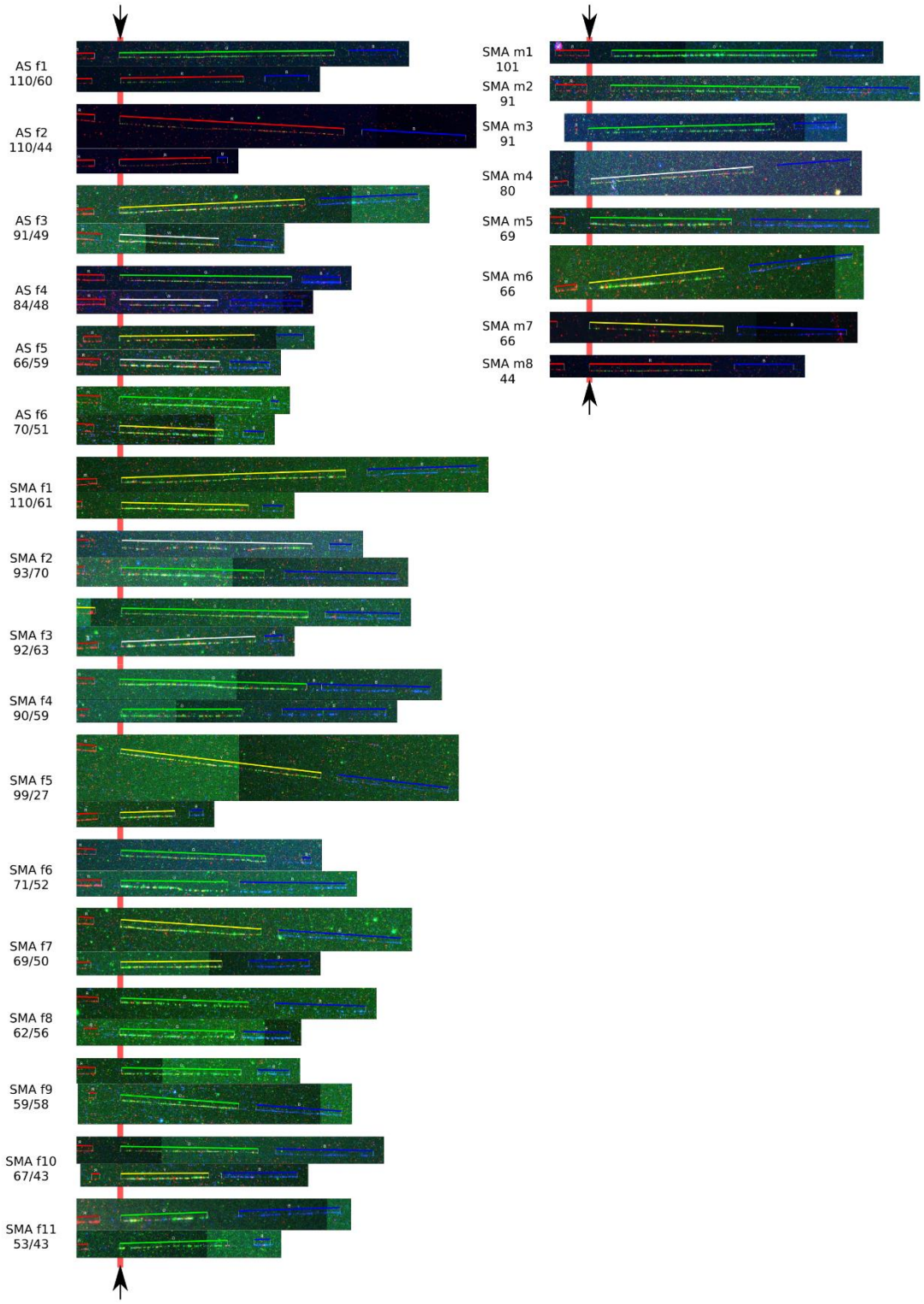


Figure 15: Results of the Molecular Combing. The copy numbers of *DXZ4* were measured by Molecular Combing in six asymptomatic females, 11 SMA-affected females and eight SMA-affected males. Given are representative pictures for each allele and the average *DXZ4* copy numbers as measured manually with the FiberCombStudio software. Samples are ordered by phenotype as female asymptomatic individuals (AS f), SMA-affected females (SMA f) and SMA-affected males (SMA m).

Next, we explored the relationship between genotypes and phenotypes. We compared the *DXZ4* copy numbers of the larger allele in female EBV cells between high and low expressers and found a significant difference ($P \approx 0.008$) (Figure 16A). The larger allele had an average copy number of 90.5 in high expressers compared to only 53.1 in low expressers. A comparison of the larger *DXZ4* allele in females and the corresponding *PLS3* levels revealed a significant linear correlation ($P \approx 0.01$, $\rho \approx 0.6$) (Figure 16B)

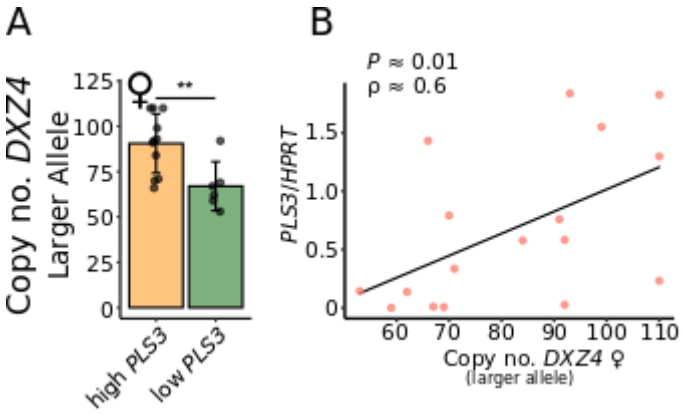


Figure 16: Relationship between the *DXZ4* copy number (larger allele) and the expression of *PLS3*. (A) Expression levels of *PLS3* were measured by qRT-PCR in 17 female EBV cell lines. The copy number of *DXZ4* was measured by Molecular Combing in the same cell lines. There is a significant difference in the *DXZ4* copy number (larger allele) between females with low (n = 6) and high (n = 11) *PLS3* expression. (B) Linear relationships between the copy number of *DXZ4* (larger allele) and *PLS3* expression in EBV cells. *HPRT* was used as housekeeping gene. Data are represented as mean ± standard deviation. The difference in the expression were analysed by Wilcoxon Rank Sum test. Linear relationships are determined by Spearman's Correlation Coefficient test (*P < 0.05; **P < 0.01; ***P < 0.001).

The smaller *DXZ4* allele in female cells did not differ between the two groups; in average high expressers had 67 and low expressers 51.5 repeats (Figure 17A). As expected, the copy number of the smaller allele did not correlate with *PLS3* levels (Figure 17B).

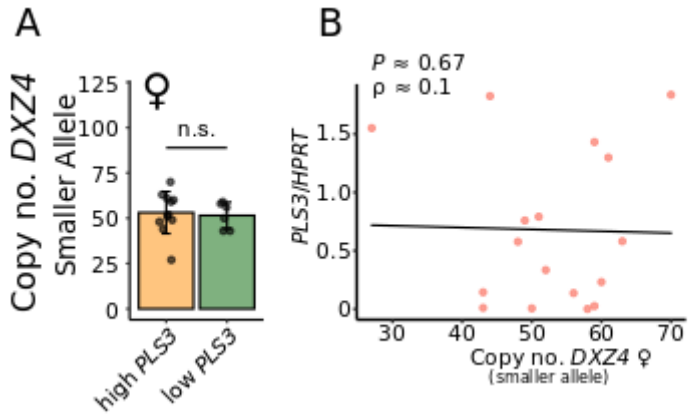


Figure 17: Relationship between the *DXZ4* copy number (smaller allele) and the expression of *PLS3*. (A) Expression levels of *PLS3* were measured by qRT-PCR in 17 female EBV cell lines. The copy number of *DXZ4* was measured by Molecular Combing in the same cell lines. (B) Linear relationships between the copy number of *DXZ4* (larger allele) and *PLS3* expression in EBV cells were not significant. *HPRT* was used as housekeeping gene. Data are represented as mean ± standard deviation. The difference in the expression were analysed by Wilcoxon Rank Sum test. Linear relationships are determined by Spearman's Correlation Coefficient test (*P < 0.05; **P < 0.01; ***P < 0.001).

In males, there was no difference in the average *DXZ4* copy number between high (79.7) and low (70) expressers (Figure 18A). Likewise, in the hemizygous males, there was no correlation between the expression of *PLS3* and the copy number of *DXZ4* (P ≈ 0.89, ρ = 0.05) (Figure 18B).

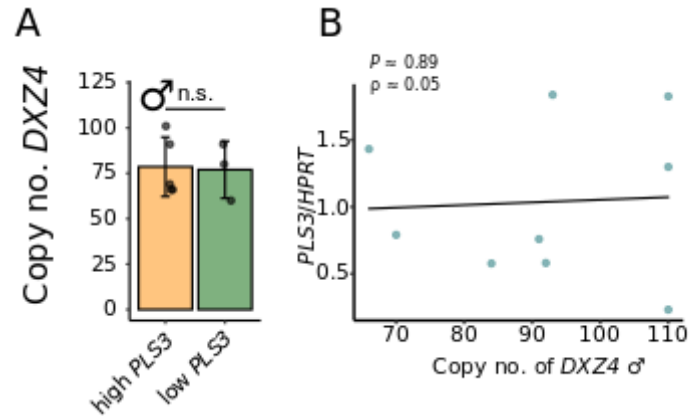


Figure 18: Relationship between the *DXZ4* copy number and the expression of *PLS3* in males. (A) Expression levels of *PLS3* were measured by qRT-PCR in eight male EBV cell lines. The copy number of *DXZ4* was measured by Molecular Combing in the same cell lines. (B) Linear relationships between the copy number of *DXZ4* and *PLS3* expression in EBV cells were not significant. *HPRT* was used as housekeeping gene. Data are represented as mean ± standard deviation. The difference in the expression were analysed by Wilcoxon Rank Sum test. Linear relationships are determined by Spearman's Correlation Coefficient test (*P < 0.05; **P < 0.01; ***P < 0.001).

Next, we compared our finding to the previous bioinformatic approach and calculated the average *DXZ4* copy number for each female sample. The largest (SMA f1) average copy number is approximately twice as big as the smallest (SMA f11). Interestingly, the average *DXZ4* copy number in females differed significantly between the two

groups ($P \approx 0.02$), showing an average of 71.8 *DXZ4* copy numbers in high expressers and only 59.3 in low expressers (Figure 19A). In line with the bioinformatic measurement of the *DXZ4* size by read depth, a direct comparison of the average *DXZ4* copy number with the *PLS3* expression levels showed a significant linear relationship ($P \approx 0.02$, $\rho \approx 0.58$) (Figure 19B and figure 12B).

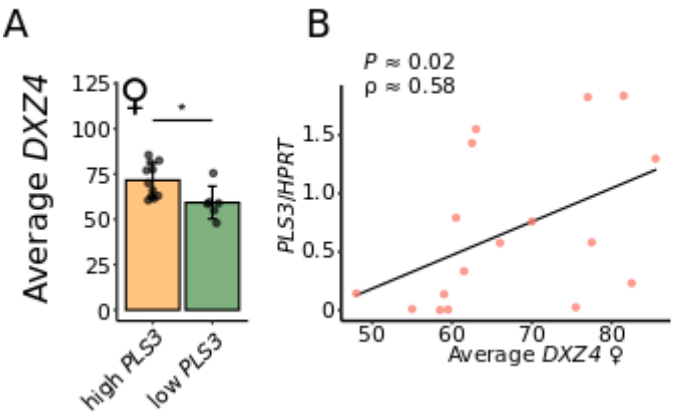


Figure 19: Average *DXZ4* copy number in *PLS3* high and low expressing EBV cells. (A) Comparison of the average *DXZ4* copy number between females with low (n = 6) and high (n = 11) *PLS3* expression. (B) Linear relationships between the copy number of *DXZ4* and *PLS3* expression in EBV cells. Comparison of the average *DXZ4* copy number measured by Molecular Combing with the expression levels of *PLS3* measured by qRT-PCR in 17 female EBV cell lines. *HPRT* was used as housekeeping gene. Data are represented as mean ± standard deviation. The difference in the expression were analysed by Wilcoxon Rank Sum test. Linear relationships are determined by Spearman's Correlation Coefficient test. (*P < 0.05; **P < 0.01; ***P < 0.001).

5.1.4. Segregation analysis of the PLS3 SNVs

The SNVs that we used to identify *PLS3* as an escape gene, were located in the exons 11 and 12 of the gene and were found in asymptomatic individuals of both families. We calculated an average SNV ratio of about 0.4, indicating that the gene is transcribed from both X-chromosomes in similar proportions and indeed escaped XCI (Figure 20A, B).

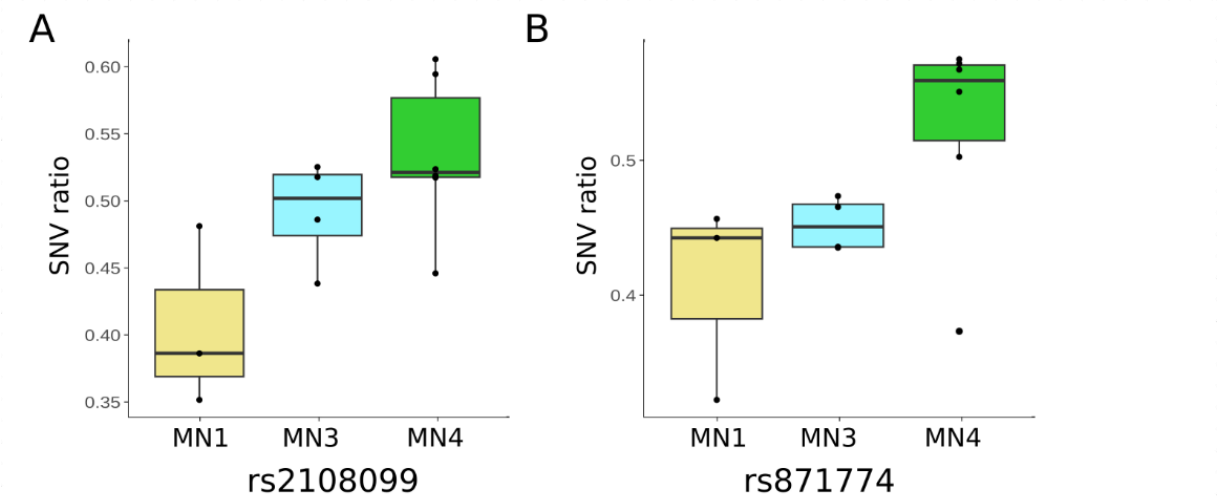


Figure 20: SNV ratios in spinal motor neurons. Given are the SNV ratios of the two biallelic exonic *PLS3* SNVs (rs2108099 and rs871774) in transcriptomes of spinal MNs of three asymptomatic females (MN1, MN3, MN4). Our data clearly show that both alleles are transcribed in equal proportions indicating that *PLS3* is escaping XCI.

The two SNVs were listed in the gnomAD data base as rs871774 (X:114,880,423-C-T (GRCh37/ hg19) and rs2108099 (X:114,879,399-C-T (GRCh37/ hg19). They were synonymous, predicted to be non-pathogenic and the total minor allele frequency was about 0.06. Interestingly, the minor allele frequency in the European population was about 0.03, while in the African populations it was 10 times higher; 0.26 for rs871774 and 0.32 for rs2108099, respectively. Furthermore, the variant rs871774 is located only 100 bp from the common intronic *PLS3* gene variant rs871773 (ChrX:114,880,523-C-T (GRCh37/ hg19)), which is associated with early colorectal cancer tumour recurrence and has been proposed to be associated with the overexpression of the *PLS3* (Szkandera et al., 2013). Linkage disequilibrium data from the National Cancer Institute website using the LDlink tool indicated a perfect linkage disequilibrium for all 3 variant combinations in the European and Asian populations, indicating that rs871774 and rs2108099 may also be associated with overexpression of *PLS3* in European populations (Machiela and Chanock, 2015). However, there is no perfect linkage disequilibrium in the African and American populations. We analysed whole-genome sequencing data of FBs of all siblings of family 1 and 2 and found that the variant rs871773 is found in all three asymptomatic females, but not in the SMA-affected brothers (Figure 21).

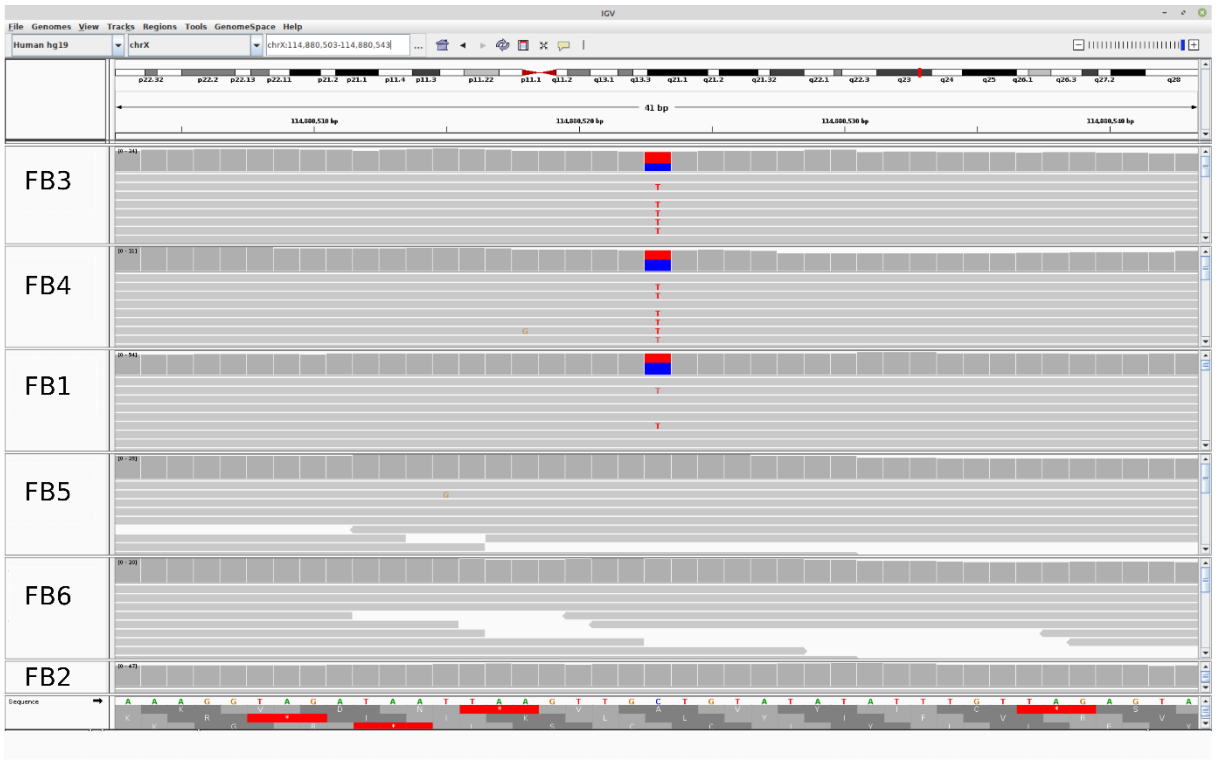
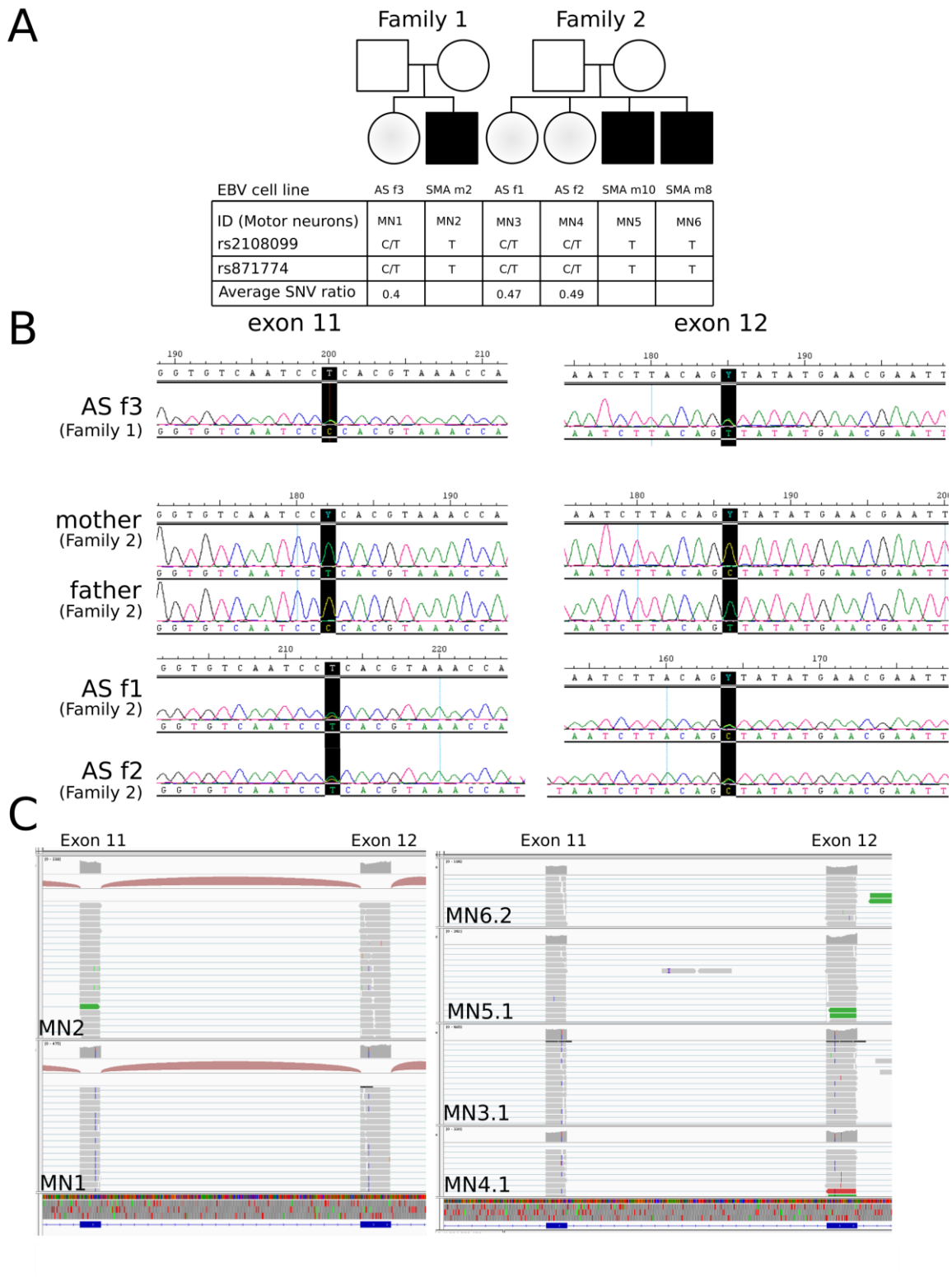


Figure 21 Detection of the variant rs871773 in whole-genome data sets. The intronic variant rs871773 is associated with early colorectal cancer recurrence and overexpression of *PLS3* in CTCs. An analysis of whole-genome sequencing data sets from FBs of both families shows that the heterozygous variant is found in all three asymptomatic females (FB1, FB3, FB4), but not in the SMA-affected brothers (FB2, FB5, FB6).

We were interested to identify other individuals that share one or both exonic *PLS3* SNVs and analysed 18 transcriptome data sets from EBV cells of female SMA patients and three healthy controls with varying degrees of *PLS3* expression. However, none of those transcriptome data sets contained any exonic *PLS3* SNVs. To segregate the two exonic *PLS3* SNVs, rs871774 and rs2108099, Sanger sequencing was performed (Figure 22A to C).



EBV cell line	AS f3	SMA m2	AS f1	AS f2	SMA m10	SMA m8
ID (Motor neurons)	MN1	MN2	MN3	MN4	MN5	MN6
rs2108099	C/T	T	C/T	C/T	T	T
rs871774	C/T	T	C/T	C/T	T	T
Average SNV ratio	0.4		0.47	0.49		

Figure 22: Biallelic exonic *PLS3* SNVs in iPSC-derived spinal motor neurons. (A) Biallelic SNVs were identified in transcriptome data of spinal MNs of both families. The female siblings were heterozygous (C/T) for both SNVs, while the male siblings carry the major (T) allele. (B) The heterozygous SNVs were validated in EBV cells by Sanger sequencing. (C) The transcriptome data show the heterozygous SNVs (C/T) in the female samples, while the male siblings carry the major T allele.

In family 1, the asymptomatic sibling is heterozygous (C/T) for both SNVs. The two *DXZ4* alleles have a copy number of 91 and 49 repeats, respectively. The SMA-

affected brother inherited the major (T) allele in both cases and the large *DXZ4* allele of 49 repeats. Since there is no recombination between the X- and Y-chromosomes outside of the pseudoautosomal regions, the minor SNV allele in the sister must be of paternal origin and in phase with the smaller *DXZ4* allele. In family 2, the two affected brothers inherited the major (T) allele for both SNVs, while two different maternal DXZ4 alleles have been identified with 60 and 45 repeats, respectively. The two sisters are heterozygous (C/T) for both SNVs. The one sister inherited one shorter *DXZ4* allele of about 60 repeats from the mother, while the other sister inherited the 45-repeat allele. The large *DXZ4* allele of about 110 repeats must be of paternal origin (Figures 14B and 15).

Overall, we have found that there is a correlation between the copy number of the larger allele and the expression levels of *PLS3* in females. This correlation has not been found for the smaller *DXZ4* allele in females, nor in males. However, we have shown by a bioinformatic approach as well as by Molecular Combing that the average *DXZ4* copy number correlates with the expression of *PLS3* in females. Increased *PLS3* levels are associated with at least one *DXZ4* allele of increased copy number (>70 repeats), while the absolute copy number between the two alleles differs significantly between high expressers and low expressers. Furthermore, all three female asymptomatic individuals carry the SNVs rs871774, rs2108099 and rs871773; the latter is assocated with early colorectal cancer tumour recurrence and has been proposed to be associated with the overexpression of *PLS3* (Szkandera et al., 2013).

5.1.5. Variable expression of PLS3 in EBV cells over time

EBV cells are blood-derived B cells, which were immortalised with Epstein-Barr Virus infection. These cells grow as non-adherent immortalised cultures and are prone to freeze-thaw cycles in the laboratory. Over the years, we measured the *PLS3* expression in these cells in various experiments. During our analysis, we encountered some rare cases, where the previously measured expression levels differed from the most current measurements. We were interested to investigate this phenomenon and screened for samples with changed expression levels. To do that, we cultivated EBV cells that were frozen from the original stock as well as cells that were frozen and thawed multiple times. By this, we identified nine samples, in which the expression of *PLS3* was strongest in the original EBV stocks, but decreased in frequently used cell

lines. All of the nine samples were previously identified *PLS3* high expressers with large *DXZ4* copy numbers (Figures 13 and 23).

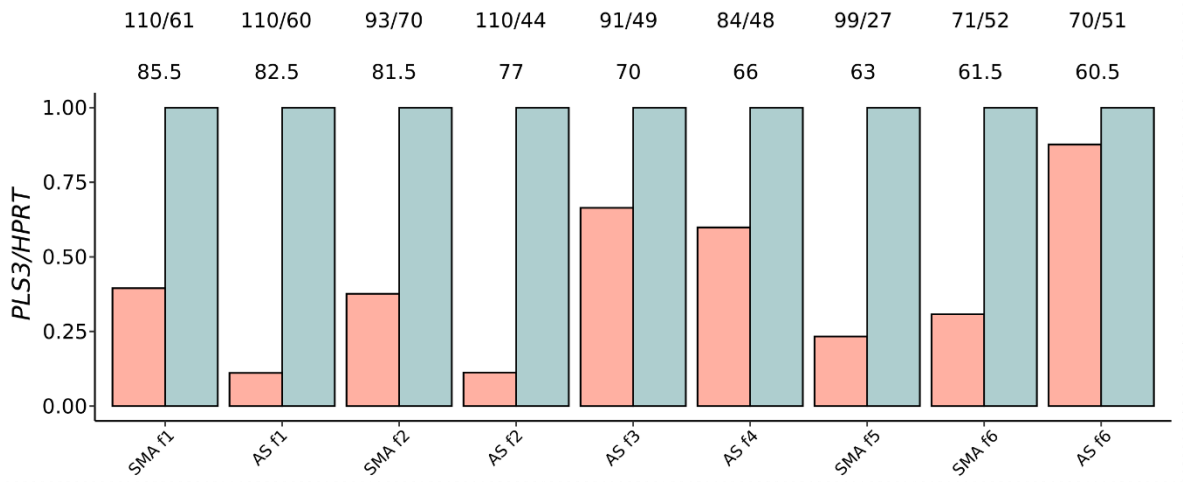


Figure 23: *PLS3* levels in nine pairs of EBV cell lines of the same origin measured by qRT-PCR. *HPRT* was used as housekeeper. The expression of *PLS3* in the frequently used sample (red) is decreased compared to the sample that was frozen from the original EBV cell line (blue). The *DXZ4* copy numbers and the absolute difference in the *DXZ4* copy number are given for each sample pair. Samples are named by phenotype as female asymptomatic individuals (AS f) and SMA-affected females (SMA f).

We designed an experiment to test different mechanisms that could explain the change of *PLS3* expression. First, we hypothesised that the XI-status has an influence on the expression levels. We proposed that over time, the balanced state of X-chromosomal inactivation shifts to a skewed state. This means that in the original culture the proportion of cells with inactivated X-chromosome of maternal and paternal origin is more or less equal, while over time the one of the two X-chromosomes is inactivated in a higher proportion, which might have an influence on the escape of *PLS3* from X-inactivation. According to our hypothesis, this would be the case, if the X-chromosome with the high copy number *DXZ4* allele is inactivated. Second, we hypothesised that even if *PLS3* escapes from XCI, there must be a transcription factor or transcriptional regulator that drives the expression of the gene. Therefore, we searched for potential transcription factors or transcription regulators that might explain the differential *PLS3* expression.

To investigate this, we extracted DNA and RNA from each sample pair and performed multiple experiments. First, the RNA was reverse transcribed into cDNA to measure the expression of *PLS3* via qRT-PCR (Figure 23). Next, we used the DNA to measure the XI-status of EBV cell lines by performing a methylation-sensitive PCR-based assay (Figure 24A) (Bertelsen et al., 2011). In this assay, genomic DNA was digested with

the methylation-sensitive restriction enzyme *Hpa*II or the non-methylation-sensitive isoshizomer *Msp*I. The digested DNA was amplified by FAM-tagged primers that target the promoter of the X-linked gene SLIT And NTRK Like Family Member 4 (*SLITRK4*, [MIM: 300562]). The targeted region harbours a restriction enzyme cutting site, which is methylated on the X_i. In the sample that was treated with *Hpa*II, the X_a is digested completely, while the X_i is unaffected by the restriction enzyme. The samples that were treated with *Msp*I are used as a digestion control and should give no signal by fluorescent capillary electrophoresis. In addition to that, one sample was treated without restriction enzyme, which means that both alleles of the *SLITRK4* promoter remain intact.

Additionally, the amplified region contains an AC-dinucleotide polymorphism. The amplified fragments were analysed by fluorescent capillary electrophoresis. The AC-dinucleotide repeat allowed for discrimination of the two alleles, while a comparison of the undigested and the *Hpa*II-digested fragments allowed us to calculate the proportions of the two alleles on the X_a (Figure 24B). The amount of each fragment was calculated as the area under the curve of the resulting chromatograms and the proportion of each allele was calculated using equation 2 (see Methods section 7.2.1.18.) (Figure 24C). A proportion of 0 to 0.4 or 0.6 to 1 of allele 1 on the active X-chromosome reflects a skewed XI-state, indicating that one of the two X-chromosomes is more often inactivated in the cell population. A balanced XI-state indicates that both alleles are active in a similar proportion.

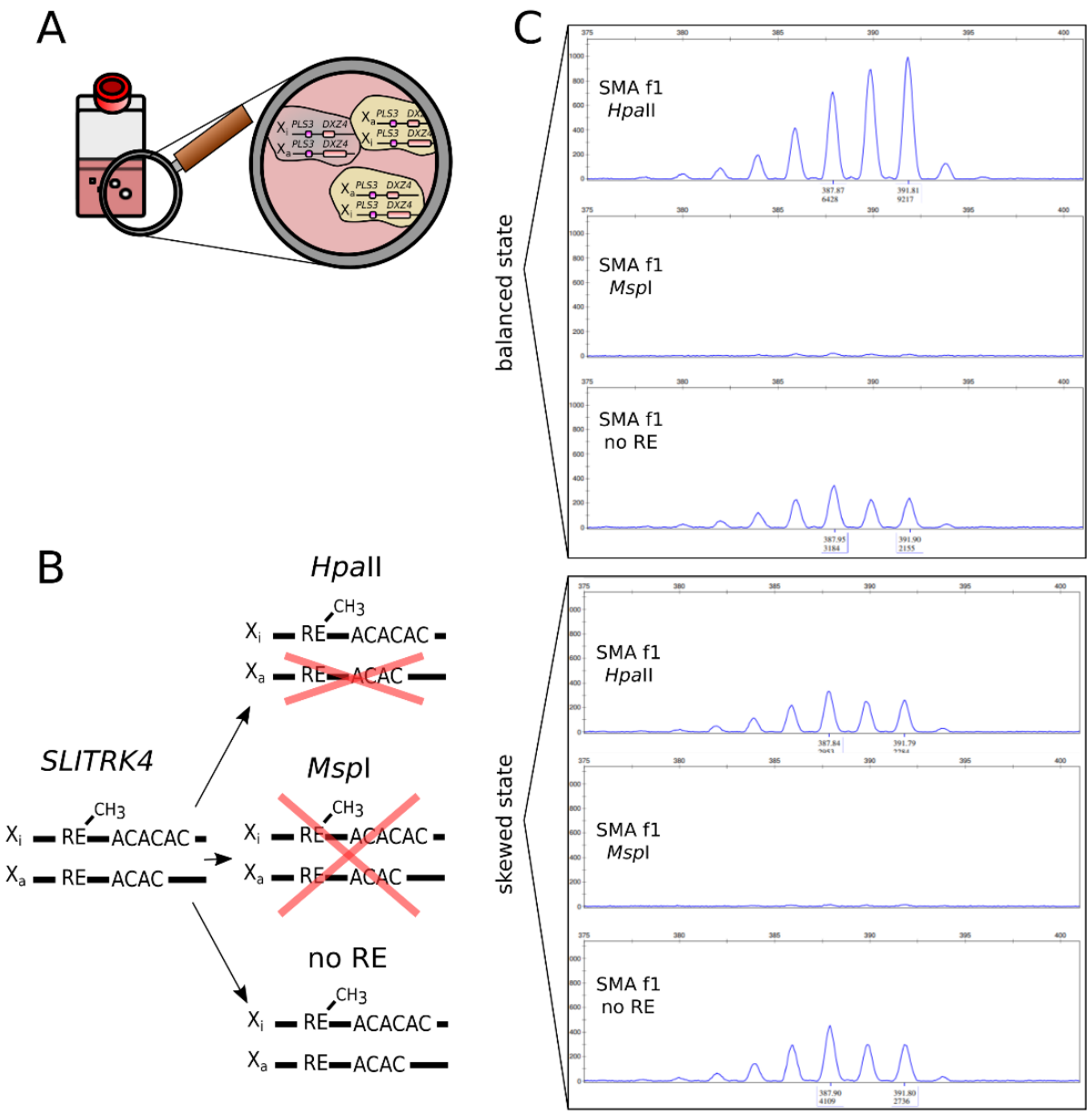


Figure 24: Methylation sensitive X-inactivation assays. (A) EBV cells grow as non-adherent cells. In each cell, either the paternal or maternal X-chromosome is inactivated, while the other chromosome is active. (B) To measure the XI-status of the population, DNA was digested with the methylation-sensitive restriction enzyme *Hpa*II or the non-sensitive isoshizomer *Msp*I. A third sample was not digested. The DNA was then amplified with FAM-tagged primers against the X-linked *SLITRK4* promoter. The target sequence harbours a restriction enzyme cutting site (methylated in X_i) and a dinucleotide polymorphism, which is used to identify the fragments by size on a fragment capillary electrophoresis. (C) Results from the fragment capillary electrophoresis. The number of curves indicates the length of the dinucleotide polymorphism, while the area under the curve indicates the amount of the fragment. By comparison of the *Hpa*II-digested and undigested curves, the XCI-status is calculated (Equation2 see methods section 7.1.1.18).

We identified three pairs of cell lines that contained a balanced XI-state in the sample that was in culture for a long period of time and skewed XI-state in the sample, which was frozen from the original EBV cell line (SMA f1, AS f1, and AS f4) (Figure 25).

However, we were not able to explain the difference in the *PLS3* expression with the changed XI-status in the other six cell lines.

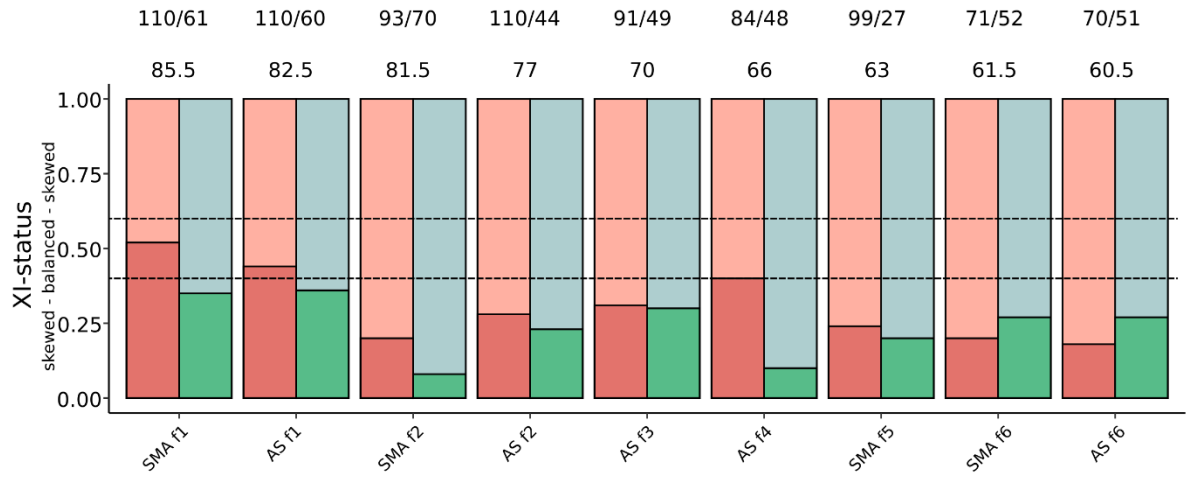


Figure 25: XI-status for each sample measured by methylation-sensitive X-inactivation assay. A balanced XI-state indicates that both alleles are active in similar proportions. A proportion of 0 to 0.4 or 0.6 to 1 of allele 1 on the active X-chromosome reflects a skewed XI-state, indicating that one of the two X-chromosomes is more often inactivated in the cell population. The cell clone that showed the higher expression of *PLS3* and was in culture for a shorter time is shown in green, while the red bar indicates the XI-status in the cell clone, which was in culture for a longer time. The *DXZ4* copy numbers and the absolute difference in the *DXZ4* copy number are given for each sample pair. Samples are named by phenotype as female asymptomatic individuals (AS f) and SMA-affected females (SMA f).

5.1.6. Identification of CHD4 as transcriptional regulator of PLS3

Our next approach was to identify putative transcription factors or transcriptional regulators that might explain the observed *PLS3* levels. To reduce biological variability caused by the *DXZ4* copy number and escape of *PLS3*, we analysed transcriptomes of 15 male EBV cell lines. We compared the four cell lines with the strongest expression of *PLS3* against eleven cell lines with lower or without *PLS3* expression and filtered for genes with a significant differential expression with a $\log_2$-fold change larger than 2 and a *P*-value smaller than 0.01 and identified 19 genes including *PLS3* (Figure 26A, Table 1).

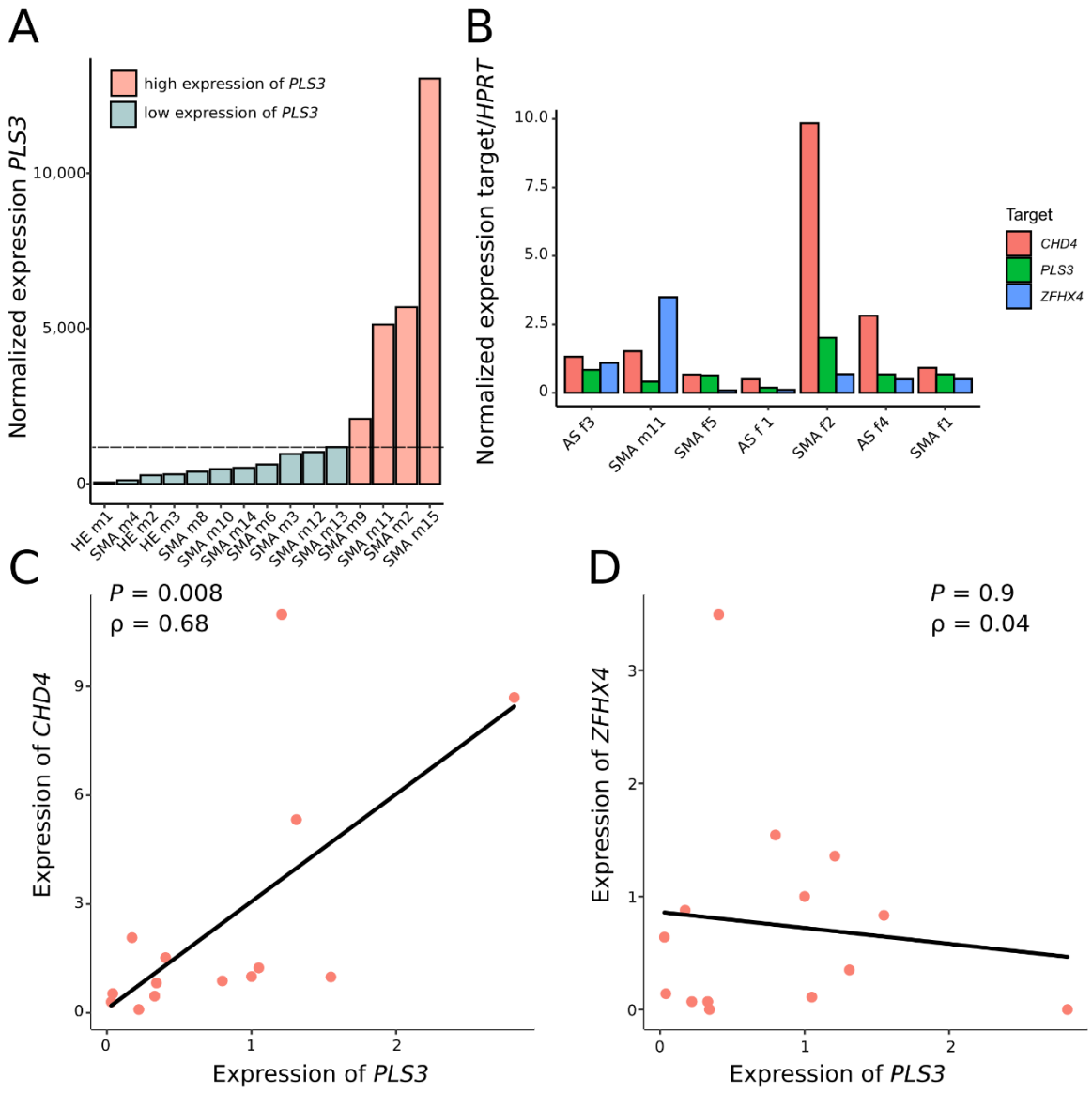


Figure 26: Transcriptional regulator *CHD4* and the transcription factor *ZFHX4* correlate with *PLS3* levels in EBV cells. (A) Differential expression analysis of male EBV cell lines with different expression of *PLS3*. (B) qRT-PCR of the genes *PLS3*, *ZFHX4* and *CHD4* in seven EBV cell lines. (C) Significant linear correlation between *CHD4* and *PLS3* expression. (D) *ZFHX4* levels compared to the expression of *PLS3* show no significant linear correlation. Linear relationships are determined by Spearman's Correlation Coefficient test. (*P < 0.05; **P < 0.01; ***P < 0.001). Samples are named by phenotype as healthy males (HE m) and SMA-affected males (SMA m).

The only gene that is involved in transcriptional regulation found in the data set was Zink Finger Homeobox 4 (*ZFHX4*, [MIM: 606940]) (Table 5). ZFHX4 is a 397 kDa transcription factor associated to several malignancies (Chudnovsky et al., 2014). The ZFHX4 protein is an interaction partner of CHD4 and this interaction regulates the glioblastoma tumour-initiating cell state (Chudnovsky, 2014). Together, ZFHX4 and CHD4 co-regulate the expression of various genes (Chudnovsky et al., 2014). We decided to further investigate these two candidates and measured the expression of

both *ZFXH4* and *CHD4* as well as *PLS3* in seven EBV cell lines by qRT-PCR and normalised the results to an internal reference (Figure 26B). We found a significant linear relationship between *CHD4* and *PLS3*, indicating a co-regulation of the two genes ($P \approx 0.009$, $\rho \approx 0.68$) (Figure 26C). *ZFHX4* was expressed only in low concentrations in EBV cells, and we did not find a linear relationship with *PLS3* (Figure 26D). For all following experiments, we concentrated solely on *CHD4* as possible candidate and measured the expression levels of the gene in all nine pairs with variable expression of *PLS3* (Figure 27).

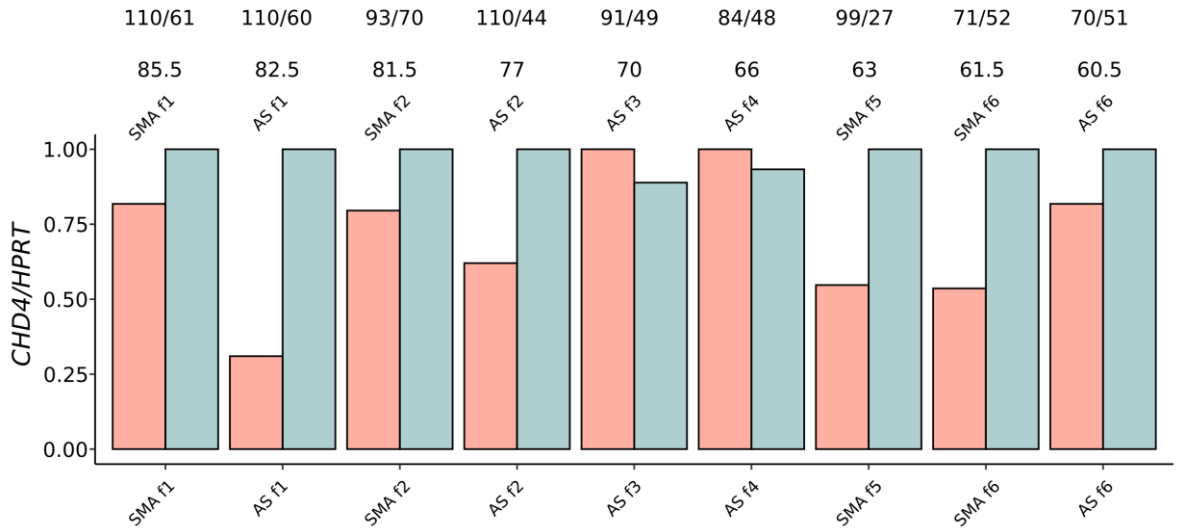


Figure 27: Expression of *CHD4* in EBV cells. *CHD4* levels in pairs of EBV cell lines of the same origin. The expression of *CHD4* in the frequently used sample (red) is changed compared to the sample that was frozen from the original EBV cell line (blue). *HPRT* was used as housekeeper.

Two of the differentially expressed escape genes, which were identified in MNs, are known to be regulated by CHD4/NuRD (Chudnovsky et al., 2014; Hoffmeister et al., 2017). Those genes were Mortality Factor 5 Like 2 (*MORF4L2*, [MIM: 300409]) and Transcription Elongation Factor A Like 4 (*TCEAL4*) (Figure 28A, B, C, D).

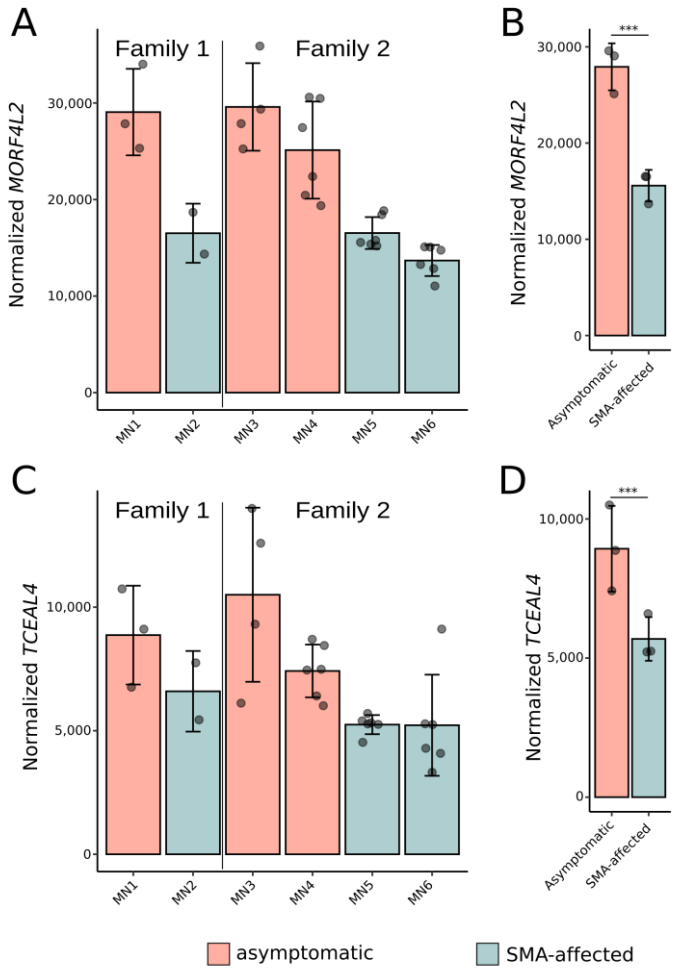


Figure 28: Expression levels of the escape genes *MORF4L2* and *TCEAL4* in iPSC-derived spinal motor neurons. (A, B) The expression of *MORF4L2* in iPSC-derived spinal MNs of both families. (C, D) The expression of *TCEAL4* in iPSC-derived spinal MNs of both families. For both genes there were significant differences in the expression between asymptomatic and SMA-affected siblings. Data are represented as mean ± standard deviation (*P < 0.05; **P < 0.01; ***P < 0.001). Asymptomatic females are shown in red, while SMA-affected male siblings are shown in green.

Both genes are localised on Xq22.2 and Xq22.1, about 11,853 and 11,953 kb apart from *PLS3* and the macrosatellite *DXZ4*, and interestingly, similar to *PLS3*, they were not differentially expressed in fibroblasts (Figure 29A, B, C, D).

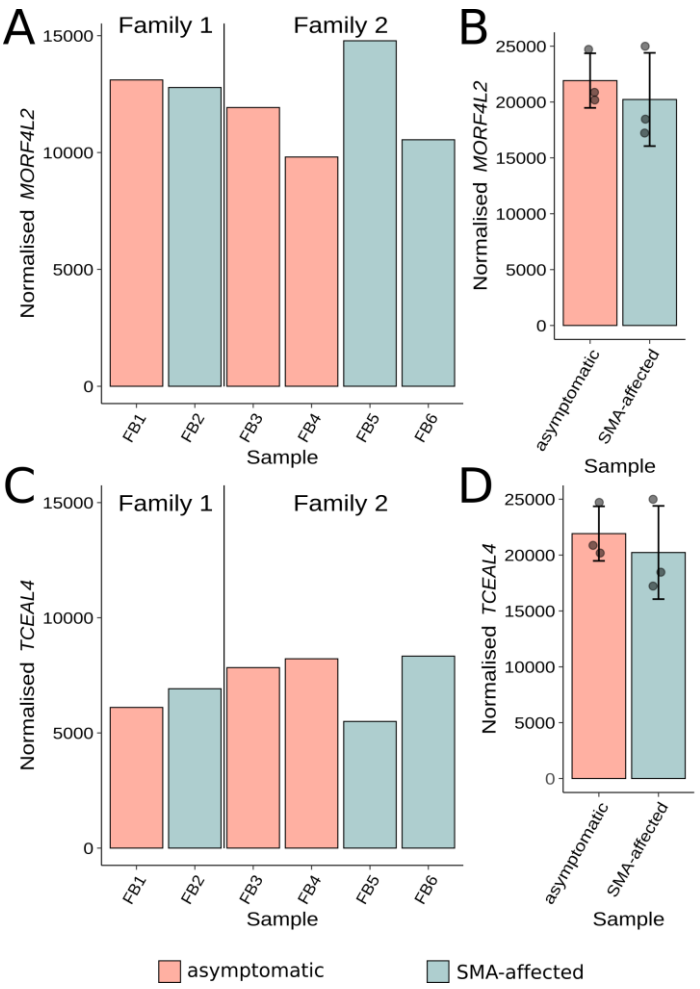


Figure 29: Expression levels of the escape genes *MORF4L2* and *TCEAL4* in FBs. (A, B) The expression of *MORF4L2* in FBs of both families. (C, D) The expression of *TCEAL4* FBs of both families. Asymptomatic females are shown in red, while SMA-affected male siblings are shown in green. Data are represented as mean ± standard deviation (*P < 0.05; **P < 0.01; ***P < 0.001).

Additionally, we investigated the lncRNA gene DXZ4-Associated Non-Coding Transcript 2 (*DANT2*, [MIM: 301004]) that is transcribed from promoters flanking *DXZ4* and is speculated to have a function in the regulation of constitutive heterochromatin formation (Figueroa et al., 2015). Overall, we found a significant linear correlation between *CHD4* and *PLS3* in EBV cells (Figure 30A).

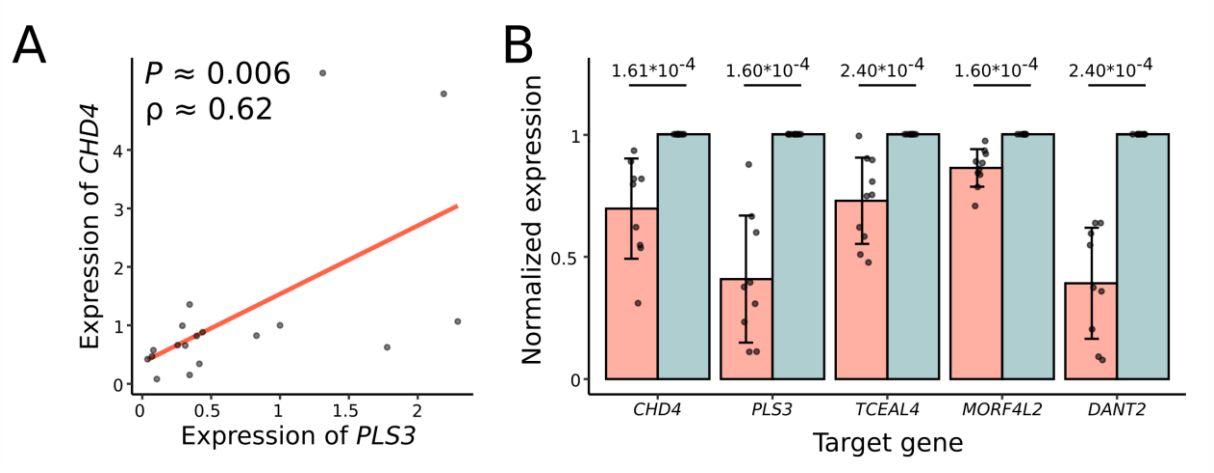


Figure 30: Comparison of the *CHD4, PLS3, TCEAL4, MORF4L2* and *DANT2* expression levels measured by qRT-PCR. (A) Linear relationships between *PLS3* and *CHD4* are determined by Spearman's Correlation Coefficient test. (B) Expression levels of *CHD4*, *PLS3*, *TCEAL4*, *MORF4L2* and *DANT2* with averages of all nine replicates. *HPRT* was used as housekeeper. The expression differences were compared by Wilcoxon rank sum test. Data are represented as mean ± standard deviation (*P < 0.05; **P < 0.01; ***P < 0.001). The expression of *CHD4* in the frequently used sample (red) is changed compared to the sample that was frozen from the original EBV cell line (blue).

We measured by qRT-PCR the expression levels of *TCEAL4*, *MORF4L2* and *DANT2* in addition to *CHD4* and *PLS3* (Figure 30B). A comparison of cells that were in culture for a longer time period changed the expression of *PLS3* significantly ($P \approx 1.61*10^{-4}$). We also found that the two escape genes *TCEAL4* ($P \approx 2.40*10^{-4}$) and *MORF4L2* ($P \approx 2.40*10^{-4}$) changed their expression levels significantly, but to a lower extent than *PLS3* (Figure 30B). Interestingly, the expression levels of *DANT2* were changed similarly to *PLS3* ($P \approx 1.60*10^{-4}$) (Figure 30B). Our data suggest a multiple influence on *PLS3* expression by the macrosatellite *DZX4*, the lncRNA *DANT2* and the transcription regulator *CHD4.*

5.1.7. Validation of CHD4 as epigenetic regulator of PLS3

The gene *CHD4* was rendered as strong candidate as epigenetic transcriptional regulator of *PLS3*. Next, we investigated the expression levels of *CHD4* in our FB and spinal MN transcriptomes (Figure 31A, B, C, D). In both cell types, there was a strong expression of *CHD4* in both asymptomatic and SMA-affected individuals.

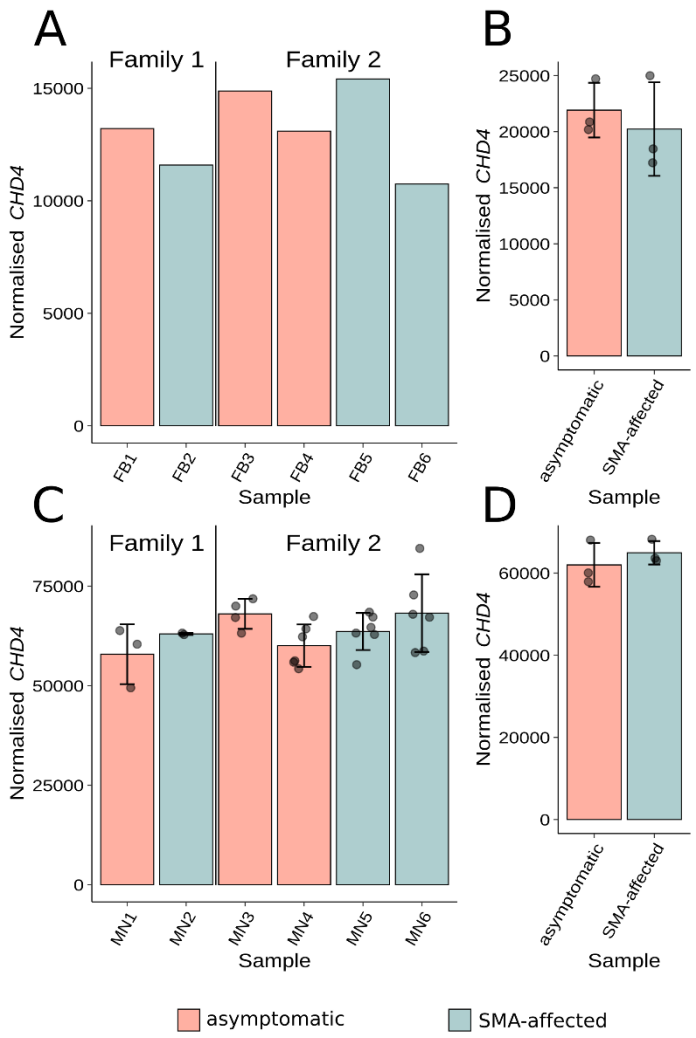


Figure 31: Expression levels of *CHD4* in FBs and MNs. (A, B) The expression of *CHD4* in FB transcriptomes of both families. (C, D) The expression of *CHD4* in spinal MN transcriptomes of both families. Asymptomatic females are shown in red, while SMA-affected male siblings are shown in green. For both cell types, *CHD4* was strongly expressed without differences between asymptomatic and SMA-affected individuals. Data are represented as mean ± standard deviation (*P < 0.05; **P < 0.01; ***P < 0.001).

We wanted to further characterise the gene and performed knock-down experiments in EBV cells using an siRNA against *CHD4*. In addition to that, we overexpressed *CHD4* in HEK293T cells by an expression plasmid. Furthermore, we performed chromatin immunoprecipitation assays followed by qRT-PCR (ChIP-qPCR) to investigate, if CHD4 directly interacts with the *PLS3* promoter. Lastly, to verify that CHD4 acts as transcriptional regulator, which activates the expression of *PLS3* directly, we performed dual-luciferase-promoter assays in HEK293T cells.

5.1.7.1. Knock-down of CHD4 in EBV cells

EBV cells were transfected with an siRNA targeting *CHD4*, or a mock control. RNA was extracted, reverse transcribed into cDNA and the expression of *CHD4*, *PLS3* and

TCEAL4 was measured via qRT-PCR. Three male and three female cell lines were used (AS f1, AS f2, AS f5, SMA m2, SMA m6, SMA m11) (Figure 32A). We obtained a knock-down efficiency of about 50% in *CHD4*, which was a significant reduction ($P < 0.01$). The *CHD4* siRNA- and mock-transfected cell originated from the same cultures and assuming that the XI-status would not change overnight, we were able to show that the downregulation of *CHD4* has a direct impact on the expression of *PLS3* without putative influences of the XI-status in female samples. The *PLS3* and *TCEAL4* levels were decreased significantly by about 40% ($P < 0.01$) (Figure 32A). We found a significant linear correlation between the *CHD4* and *PLS3* expression levels ($P \approx 5.9*10^{-8}$, $\rho \approx 0.978$) (Figure 32B).

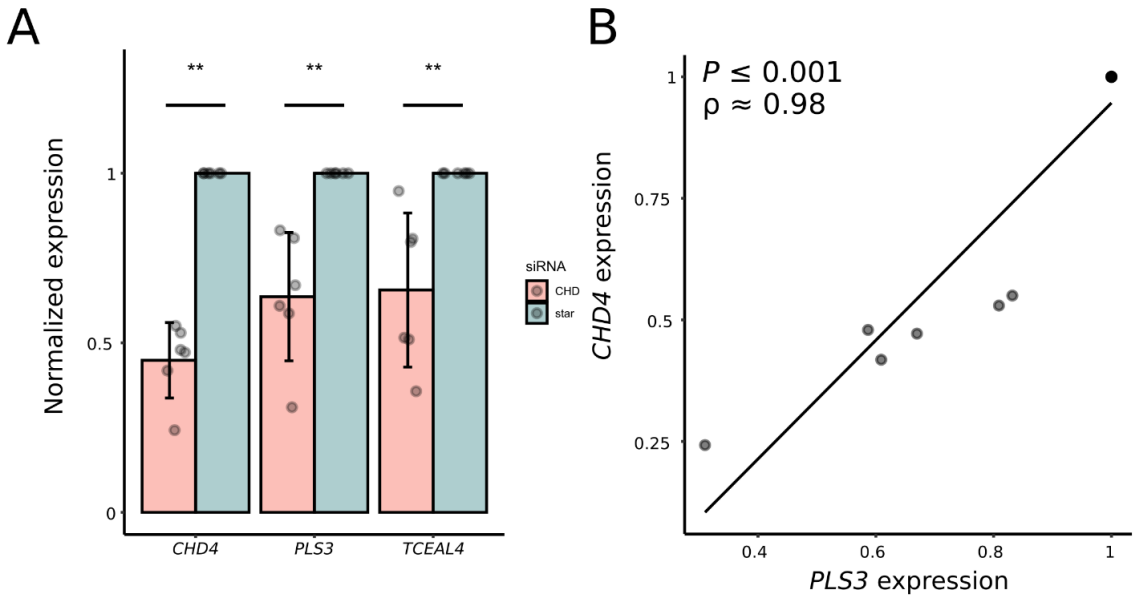


Figure 32: CHD4 is a transcriptional regulator of *PLS3*. (A) siRNA-mediated knock-down of *CHD4* in EBV cells against a mock control. The differences in the expression were compared by Wilcoxon Rank Sum tests. Data are represented as mean ± standard deviation. Knock-down of *CHD4* leads to significant reduction of *PLS3* and *TCEAL4* expression levels. (B) Comparison of the *CHD4* and *PLS3* expression levels. A linear relationship was determined by Spearman's Correlation Coefficient test indicating that both genes are co-regulated in EBV cells (*P < 0.05; **P < 0.01; ***P < 0.001).

5.1.7.2. Overexpression of CHD4 in HEK293T cells

We first established an overexpression vector that expresses CHD4 under a Cytomegalovirus (*CMV*) promoter. Next, we transfected 125,000 HEK293T cells with increasing concentrations of the CHD4 expressing plasmid DNA (25, 50, 75 and 100 ng), which was verified by qRT-PCR (Figure 33A). The subsequent qRT-PCR analysis of *PLS3* expression showed a significant increase of *PLS3* levels: 75 ng ($P \approx 0.03$) and 100 ng ($P \approx 0.03$) (Figure 33B). Overall, we found a significant linear correlation

between the *CHD4* and *PLS3* expression levels in HEK293T cells ($P \approx 1.7*10^{-3}$, $\rho \approx 0.65$) (Figure 33C).

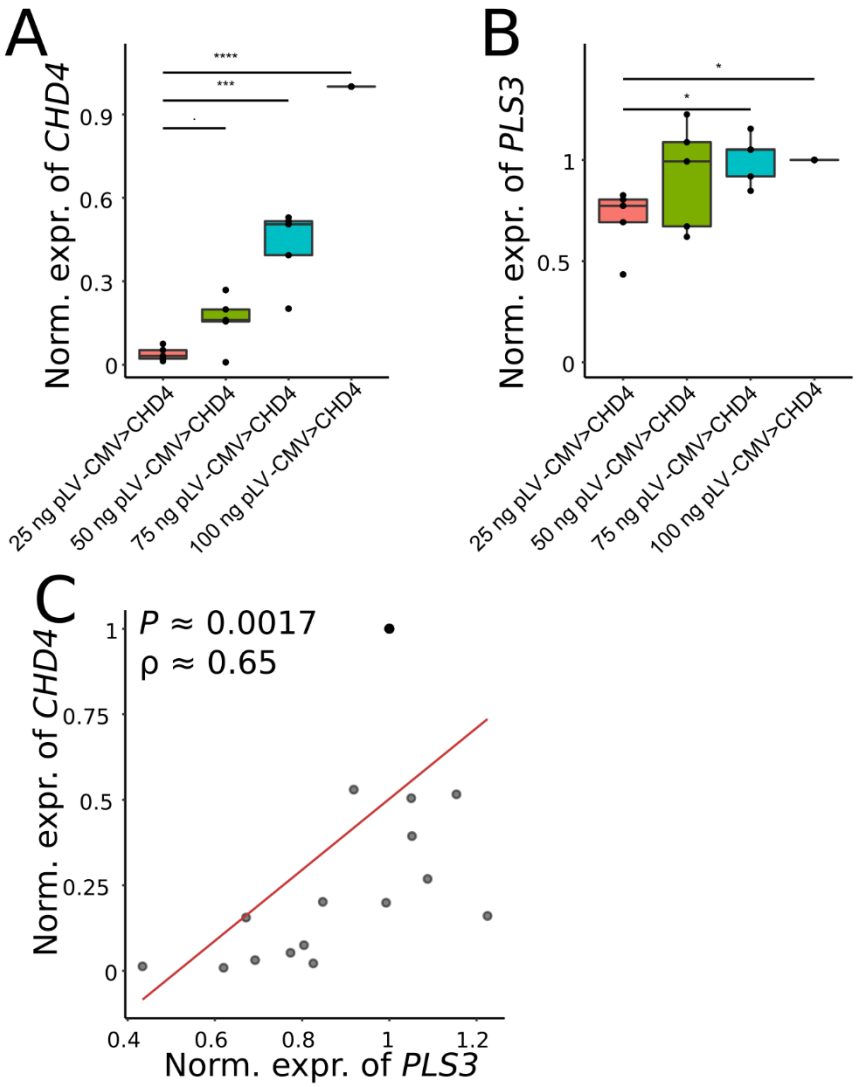


Figure 33: Overexpression of CHD4 associates with upregulation of *PLS3* in HEK293T cells. (A, B) qRT-PCR of 125,000 HEK293T cells treated with different amounts of a CHD4 overexpression vector. Comparison of the *CHD4* and *PLS3* expression levels in cells transfected with different amounts of a CHD4 expression vector. (C) Linear relationships are determined by Spearman's correlation coefficient test (*P < 0.05; **P < 0.01; ***P < 0.001).

However, we found a large variability in the expression of *PLS3* at these low concentrations of the overexpression vector. Therefore, we repeated the experiment with higher *CHD4* plasmid concentrations and transfected 450,000 HEK293T cells with 250 ng, 500 ng, 1000 ng, and 1500 ng, DNA respectively. We found significant differences in the expression of *CHD4* between the lowest vector concentration and the cells that were treated with 1000 ng ($P \approx 4.91*10^{-3}$), as well as 1500 ng of the vector ($P \approx 9.1*10^{-4}$) (Figure 34A). Furthermore, we found significant differences in the

expression of *PLS3* between the lowest concentration and the cells that were treated with 1000 ng ($P \approx 5.51*10^{-3}$), as well as 1500 ng of the vector ($P \approx 3.3*10^{-4}$) (Figure 34B). A comparison of the *CHD4* and the *PLS3* expression revealed a strong significant linear correlation ($P \approx 1.4*10^{-3}$, $\rho \approx 0.8$) (Figure 34C). Our data clearly show that the co-regulation between *CHD4* and *PLS3* can be explained by a direct influence of CHD4 on *PLS3* expression.

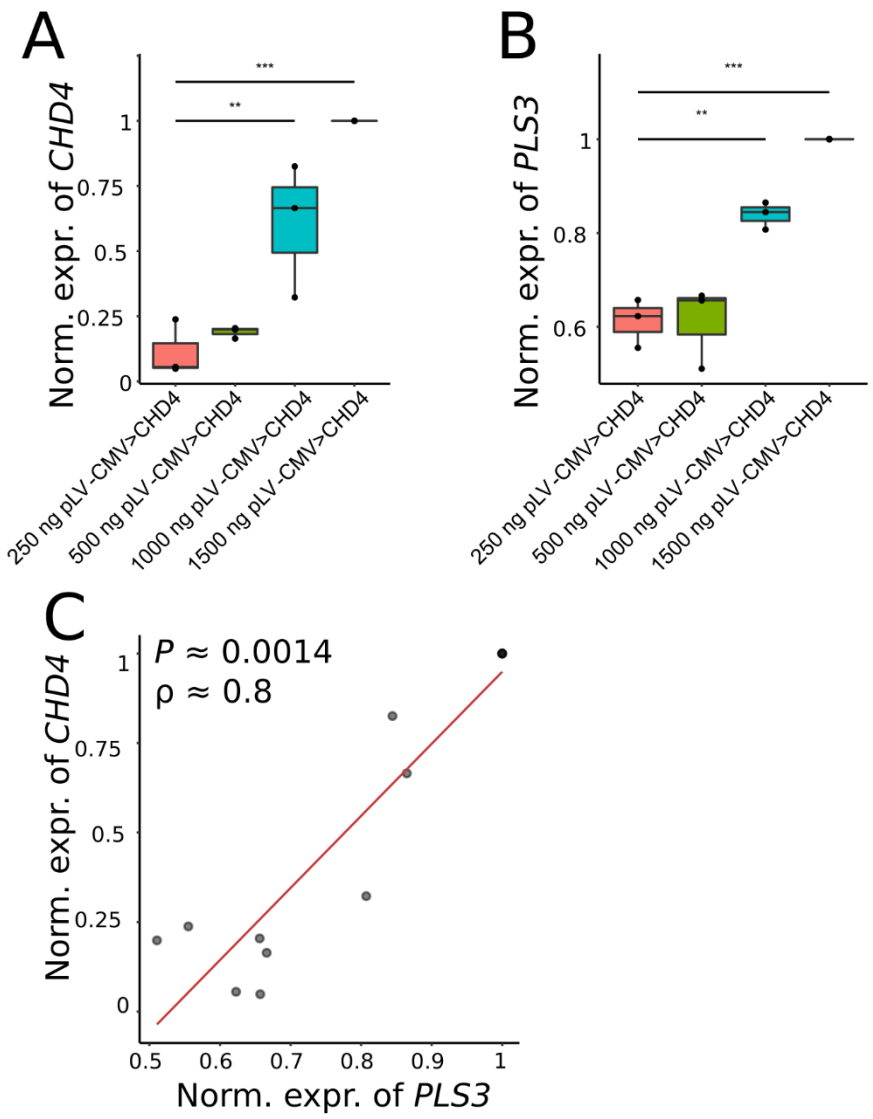


Figure 34: Overexpression of CHD4 associates with upregulation of *PLS3* in HEK293T cells. (A, B) qRT-PCR of 450,000 HEK293T cells treated with different amounts of a CHD4 expression vector. Comparison of the *CHD4* and *PLS3* expression levels in cells transfected with different amounts of a CHD4 expression vector. (C) Linear relationships are determined by Spearman's correlation coefficient test. (*P < 0.05; **P < 0.01; ***P < 0.001).

5.1.7.3. Chromatin immunoprecipitation of CHD4

CHD4/NuRD either activates or represses transcription depending on the cellular context (Hoffmeister et al., 2017). It is under debate, if CHD4 has DNA binding capacity

outside of the NuRD complex (Arends et al., 2019). We aimed to investigate this by applying ChIP-qPCR in two EBV cell lines (SMA f5, AS f2). DNA from the EBV cell lines was sheared and hybridised with a ChIP-grade CHD4 antibody. The immunoprecipitated DNA was amplified by RT-PCR using primers directed to the promoters of *PLS3* and *TCEAL4* as well as Myoglobin (*MB*, [MIM: 160000]) as negative control (see methods section 7.1.1.). The CHD4 antibody preferentially bound to both *PLS3* and *TCEAL4*, but not the *MB* control (Figure 35A, B). Furthermore, we hybridised DNA with an antibody against Immunoglobulin Heavy Constant Gamma 1 (IgG) to control for specificity of the immunoprecipitation. The IgG antibody bound the *PLS3* and *TCEAL4* promoter regions to a low degree validating the specificity of the CHD4 antibody. Finally, the DNA was hybridised with an anti-CTCF antibody, which preferentially bound to imprinted maternally expressed noncoding transcript (*H19*, [MIM: 103280]) DNA (positive control) and failed to bind *MB* DNA (negative control).

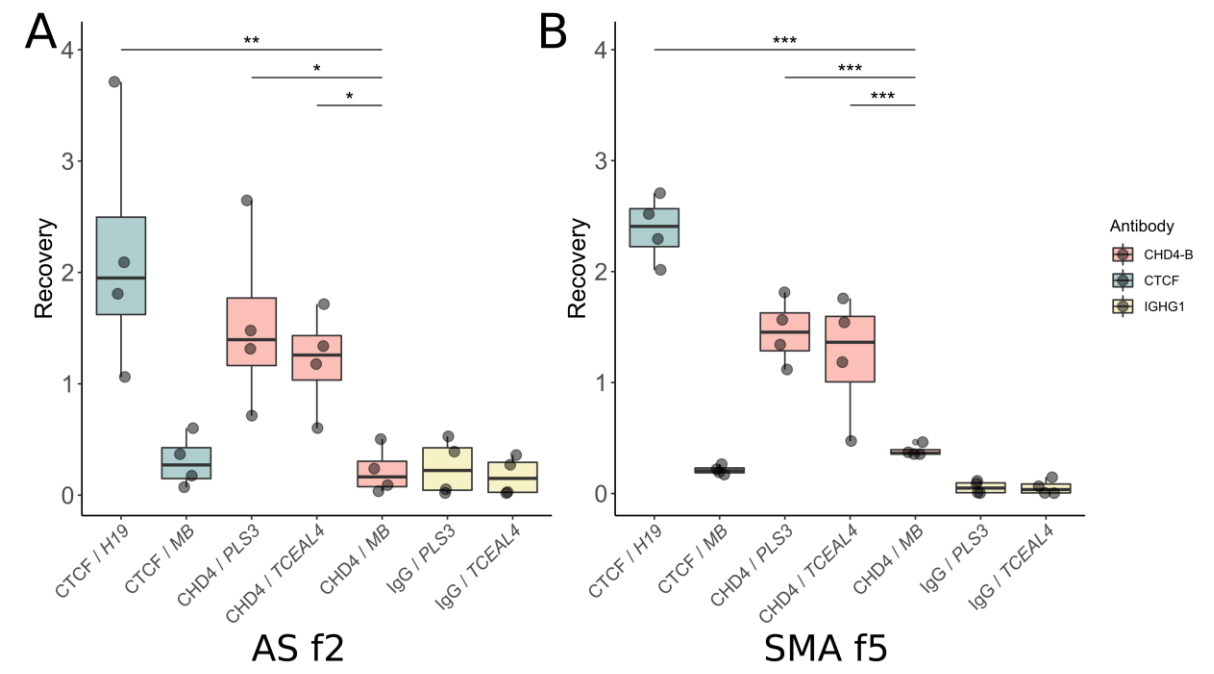


Figure 35: CHD4 directly interacts with the promoter of *PLS3*. (A, B) Chromatin immunoprecipitation of CHD4 in two different EBV cell lines. The CHD4 antibody preferentially bound to both *PLS3* and *TCEAL4* but not to *MB* (negative control). The mean ± standard deviation for n = 3 independent replicates is given. The variances were analysed by an ANOVA followed by *post-hoc* Dunnett's test for comparing several treatments with one control. Data are represented as mean ± standard deviation (*P < 0.05; **P < 0.01; ***P < 0.001).

In the sample AS f2, we found significant differences between the negative control (CHD4 / *MB*) and the positive control (CTCF / *H19*) ($P \approx 1.2*10^{-2}$). The negative control differed significantly from the target (CHD4 / *PLS3*) ($P \approx 0.015$) and (CHD4 / *PLS3*) (*P*

≈ 0.031) (Figure 35A). In the sample SMA f5, we found significant differences between the negative control (CHD4 / *MB*) and the positive control (CTCF / *H19*) ($P \approx 8.6*10^{-9}$). The negative control differed significantly from the target (CHD4 / *PLS3*) ($P \approx 6.8*10^{-4}$) and (CHD4 / *PLS3*) ($P \approx 6.3*10^{-5}$) (Figure 35B). Our data set supports the hypothesis that CHD4 is indeed able to bind DNA specifically outside of the NuRD complex and the promoters of both *PLS3* and *TCEAL4* are targets of CHD4.

5.1.4.4. Dual-luciferase promoter assays

The CHD4/NuRD complex is most often associated with repression, rather than activation of gene expression. It was suggested, that "activation of transcription" by the complex is achieved by promoter occupancy (Yamada et al., 2014). By this mechanism, CHD4/NuRD blocks the promoter region of a target gene, leading to a repression. The activation is then received by the usage of an alternative promoter. Other groups stated that CHD4/NuRD interacts with additional specific transcription factors, such as ZFXH4, to directly regulate gene expression of target genes (Figure 36) (Chudnovsky, 2014).

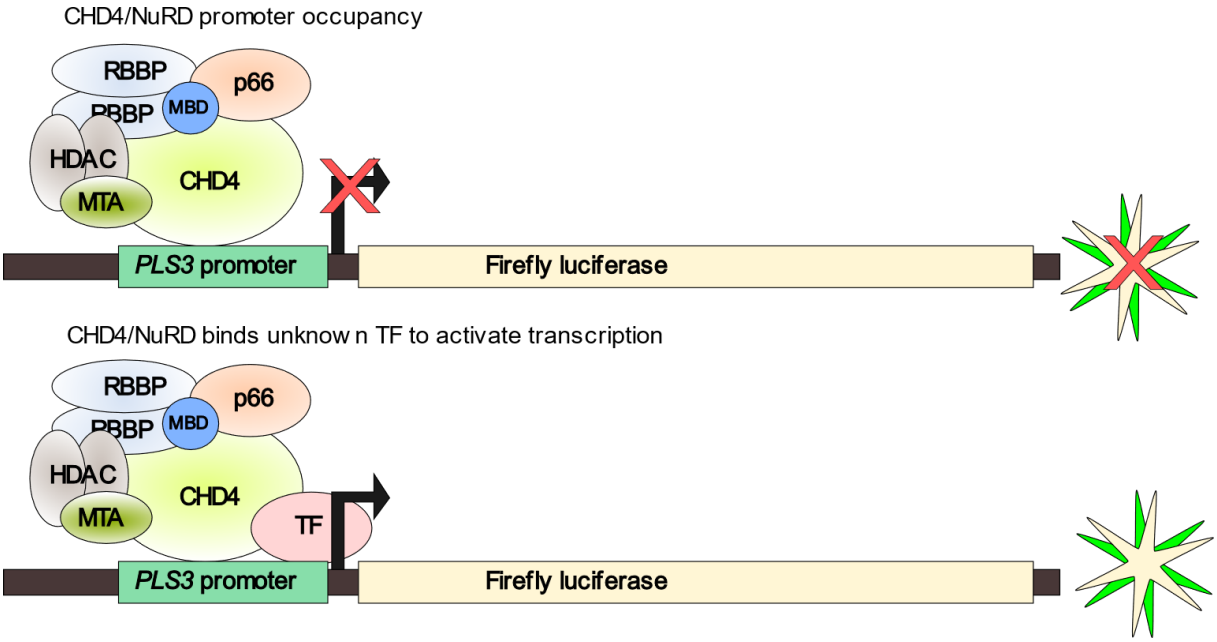


Figure 36: Direct interaction of CHD4 with the promoter of *PLS3*. If CHD4/NuRD regulates the expression of *PLS3* by promoter occupation (upper part), a dual-luciferase promoter assay would generate no luciferse signal. If CHD4/NuRD is able to interact directly with the *PLS3* promoter (lower part), the reporter gene should be expressed and a luciferase signal should be detectable.

To investigate if CHD4 is able to directly interact with the promoter of *PLS3*, we performed dual-luciferase promoter assays in 125,000 HEK293T cells. The cells in all treatment groups and the controls were co-transfected with 0.25 ng of a Renilla-

luciferase reporter construct as internal reference. In addition to that, the treatment groups were transfected with 20 ng Firefly-luciferase vector *PLS3* promoter (1022 bp) and sequential concentrations (0 ng, 6.25 ng, 12.5 ng, and 25 ng) of the *CHD4* overexpression vector, which was previously established in the overexpression assays. Furthermore, we included a positive control, which was co-transfected with 50 ng of a Firefly-luciferase vector under a *CMV* promoter. One negative control was co-transfected with 10 ng of a Firefly-luciferase vector without promoter (empty control). A second negative control was transfected only with 0.25 ng of the Renilla-luciferase vector as internal reference. We compared the endogenous *CHD4* expression (0 ng CHD4 overexpression vector) to the three treatment groups. The relative promoter activity was significantly increased by about 50% after transfection with 25 ng CHD4 overexpression vector (Figure 37A). Generally, we obtained an increase of the relative *PLS3* promoter activity with rising concentrations of *CHD4* and found a significant linear correlation ($P \approx 0.001$, $\rho \approx 0.89$) (Figure 37B).

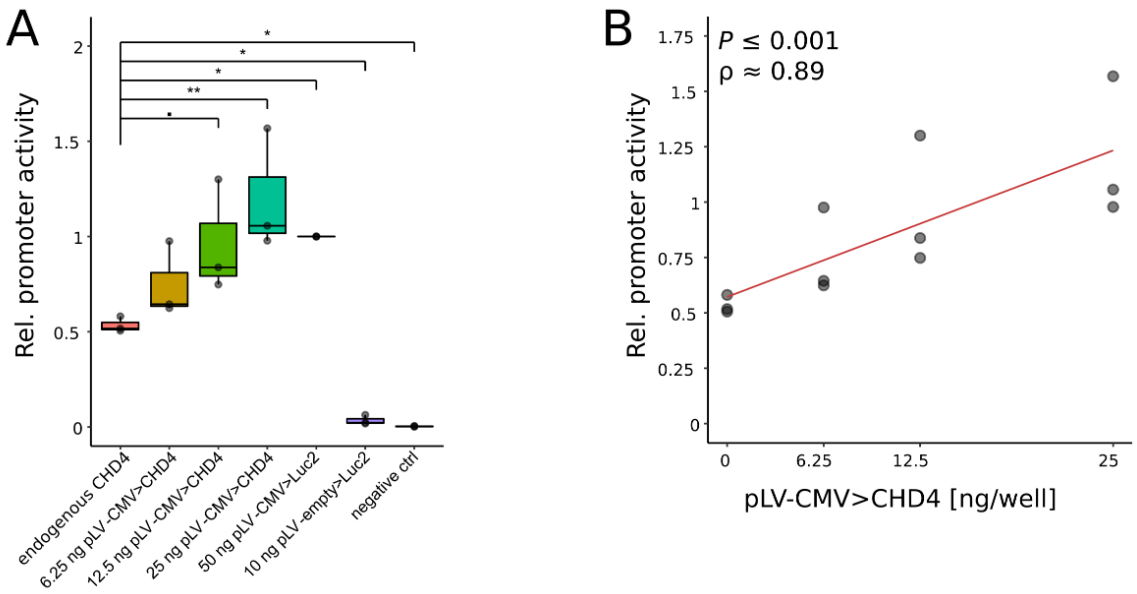


Figure 37: Dual-luciferase-promoter assays in EBV cells. (A) Activity of the *PLS3* promoter in a dual-luciferase-promoter assay. HEK293T cells were transfected with multiple concentrations of a CHD4 expression vector. The mean ± standard deviation for n = 3 independent replicates is given. The variances were analysed by an ANOVA followed by a *post-hoc* Dunnett's test for comparing several treatments with one control. (B) A direct comparison of the relative promoter activity with the amount of *CHD4*. Linear relationships are determined by Spearman's correlation coefficient test deviation (*P < 0.05; **P < 0.01; ***P < 0.001).

Our results indicate that CHD4 is directly interacting with the *PLS3* promoter. This interaction positively regulates the expression of *PLS3*. Overall, we have found that

CHD4 is an epigenetic regulator for *PLS3*, which is able to positively regulate the expression of the target gene by direct interaction.

5.1.7.4. CHD4 and the rs1639122

Lastly, we analysed our transcriptome data sets for *CHD4* variants that might be associated with the expression of the gene.

The variant *CHD4* rs1639122 (Chr12: 6,711,147-C-A (GRCh37/ hg19)) is associated with bone mineral density in a GWAS study (Liu et al., 2020). It has an allele frequency of about 0.4 according to the gnomAD database. We analysed variants in the MN, FB and EBV transcriptomes and found that this variant in 20 data sets (Table 6). The variant is found in four asymptomatic individuals (AS f1 (T/T), AS f2 (T/T), AS f5 (C/T), and AS f6 (C/T)). Furthermore, the variant was found in seven SMA-affected females (SMA f3 (C/T), SMA f6 (T/T), SMA f7 (T/T), SMA f9 (C/T), SMA f10 (T/T), SMA f12 (C/T), and SMA f15 (C/T)), in seven SMA-affected males (SMA m2 (C/T), SMA m6 (C/T), SMA m8 (C/T), SMA m10 (C/T), SMA m11 (C/T), SMA m15 (C/T), SMA m15 (C/T)), and in two healthy males (HE m2 (C/T), and HE m3 (C/T)). Of those 20 samples, eight samples were known *PLS3* high expressers (measured by qRT-PCR). For another eight samples, we have no data from the qRT-PCR and four samples were low expressers (Figure 13, Table 6). We conclude that the *CHD4* variant rs1639122 has no significant effect on the expression of *PLS3* in EBV cells.

5.2. Part II: The impact of PLS3 on biomarkers of SMA

5.2.1. Experimental design

We aimed to investigate the impact of *PLS3* overexpression on the plasma concentration of the SMN protein and on six putative plasma biomarker proteins (COMP, DPP4, SPP1, CLEC3B, VTN, and AHSG) that were previously identified by the BforSMA study and verified in the severely affected Δ7SMA mouse model (Arnold et al., 2016; Finkel et al., 2012; Kobayashi et al., 2013). We generated multiple genotypes that reflect different disease severities. The murine *Smn* was knocked out completely in "SMA" animals, while a human transgene was introduced (*Smn*$^{KO/KO}$; *SMN2*$^{tg/0}$). "HET" animals were heterozygous for the murine *Smn* (*Smn*$^{KO/WT}$; *SMN2*$^{tg/0}$). We crossed in *PLS3* overexpressing human transgenic animals and obtained "SMA-*PLS3*het" (*Smn*$^{KO/KO}$; *SMN2*$^{tg/0}$; *PLS3*$^{tg/0}$) and "SMA-*PLS3*hom" (*Smn*$^{KO/KO}$; *SMN2*$^{tg/0}$; *PLS3*$^{tg/tg}$) mice as well as "HET-*PLS3*het" (*Smn*$^{KO/WT}$; *SMN2*$^{tg/0}$; *PLS3*$^{tg/0}$) and "HET-*PLS3*hom" (*Smn*$^{KO/WT}$; *SMN2*$^{tg/0}$; *PLS3*$^{tg/tg}$) mice. In addition, WT C57BL/6N mice were used as controls. Each group consisted of three males and three females. As these untreated mice generally have a mean survival or about 16 days with symptoms occurring at P5 on a C57BL/6N background, we collected blood samples of all mice at P10 (Figure 38) (Ackermann et al., 2013; Hosseinibarkooie et al., 2016). As the protective effect of genetic modifiers in SMA takes place in mild SMA, we treated the same genotypes with a low-dose *SMN*-ASO (30 µg at P2 and P3) pre-symptomatically by subcutaneous injections (Hosseinibarkooie et al., 2016; Oprea et al., 2008; Riessland et al., 2017). Since low-dose *SMN*-ASO-treated SMA mice survived approximately 26 days (Hosseinibarkooie et al., 2016), we collected blood samples at P10 and P21, which further allowed a longitudinal analysis. The analysis of SMN plasma protein concentration was carried out by PharmOptima using a SMN ECL immunoassay as previously described (Zaworski et al., 2016). The six plasma biomarkers were analysed using a 10-plex assay panel developed by PharmOptima using commercially available antibody and calibrator reagents and multiplex assay plates used in this study were custom manufactured by Meso Scale Discovery (Arnold et al., 2016; Kobayashi et al., 2013).

Thus, we investigated SMN and the six plasma proteins in all seven ASO-treated groups at two different time points (21 groups; Figure 38). Overall, we investigated the

plasma concentration of SMN, COMP, DPP4, SPP1, CLEC3B, VTN and AHSG 21 different groups (Table 7).

Group name	murine *Smn*	human SMN2 transgene	human PLS3 transgene	ASO-treatment	Sampling day	Pooled
SMA	$Smn^{ko/ko}$	$SMN2^{tg/0}$	w/o	no	P10	SMA P10
SMA:*PLS3*het	$Smn^{ko/ko}$	$SMN2^{tg/0}$	$PLS3^{tg/0}$	no	P10	SMA P10
SMA:*PLS3hom*	$Smn^{ko/ko}$	$SMN2^{tg/0}$	$PLS3^{tg/tg}$	no	P10	SMA P10
HET	$Smn^{ko/wt}$	$SMN2^{tg/0}$	w/o	no	P10	HET P10
HET:*PLS3*het	$Smn^{ko/wt}$	$SMN2^{tg/0}$	$PLS3^{tg/0}$	no	P10	HET P10
HET:*PLS3hom*	$Smn^{ko/wt}$	$SMN2^{tg/0}$	$PLS3^{tg/tg}$	no	P10	HET P10
WT	$Smn^{wt/wt}$	w/o	w/o	no	P10	WT P10
SMA+ASO	$Smn^{ko/ko}$	$SMN2^{tg/0}$	w/o	yes	P10	SMA+ASO P10
SMA:*PLS3*het+ASO	$Smn^{ko/ko}$	$SMN2^{tg/0}$	$PLS3^{tg/0}$	yes	P10	SMA+ASO P10
SMA:*PLS3hom*+ASO	$Smn^{ko/ko}$	$SMN2^{tg/0}$	$PLS3^{tg/tg}$	yes	P10	SMA+ASO P10
HET+ASO	$Smn^{ko/wt}$	$SMN2^{tg/0}$	w/o	yes	P10	HET+ASO P10
HET:*PLS3*het+ASO	$Smn^{ko/wt}$	$SMN2^{tg/0}$	$PLS3^{tg/0}$	yes	P10	HET+ASO P10
HET:*PLS3hom+ASO*	$Smn^{ko/wt}$	$SMN2^{tg/0}$	$PLS3^{tg/tg}$	yes	P10	HET+ASO P10
WT+ASO	$Smn^{wt/wt}$	w/o	w/o	yes	P10	WT+ASO P10
SMA+ASO	$Smn^{ko/ko}$	$SMN2^{tg/0}$	w/o	yes	P21	SMA+ASO P21
SMA:*PLS3*het+ASO	$Smn^{ko/ko}$	$SMN2^{tg/0}$	$PLS3^{tg/0}$	yes	P21	SMA+ASO P21
SMA:*PLS3hom*+ASO	$Smn^{ko/ko}$	$SMN2^{tg/0}$	$PLS3^{tg/tg}$	yes	P21	SMA+ASO P21
HET+ASO	$Smn^{ko/wt}$	$SMN2^{tg/0}$	w/o	yes	P21	HET+ASO P21
HET:*PLS3*het+ASO	$Smn^{ko/wt}$	$SMN2^{tg/0}$	$PLS3^{tg/0}$	yes	P21	HET+ASO P21
HET:*PLS3hom+ASO*	$Smn^{ko/wt}$	$SMN2^{tg/0}$	$PLS3^{tg/tg}$	yes	P21	HET+ASO P21
WT+ASO	$Smn^{wt/wt}$	w/o	w/o	yes	P21	WT+ASO P21

Figure 38: Overview of the 21 sample groups. Given are the group name, the genotype of the murine *Smn*, the genotype of the human *SMN2* transgene, the genotype of the human *PLS3* transgene, if the group was treated with a *SMN* ASO, the sampling day and the name of the pooled group. Each group consisted of three male and three female mice. The plasma concentration of SMN, COMP, SPP1, DPP4, CLEC3B, AHSG and VTN was measured in each mouse. SMA mice (SMA, SMA:*PLS3*het, and SMA:*PLS3*hom) are shown in green, HET mice (HET, HET:*PLS3*het, and HET:*PLS3*hom) are shown in red, and wild-type control mice (WT) are shown in yellow.

Before we began with the analysis of the biomarker data, we measured the expression of murine Pls3 in brain and muscle tissue lysates as well as in blood of WT mice. Endogenous *Pls3* was not expressed in blood but present in brain and muscle tissues (Figure 39A). Furthermore, we measured the expression of murine *Pls3* and the human transgene *PLS3* in three male and three female transgene mice. The *PLS3* transgene was expressed in blood of transgenic mice in a low amount (Figure 39B). There was no sex-specific difference in the expression between male and female mice (Figure 39C).

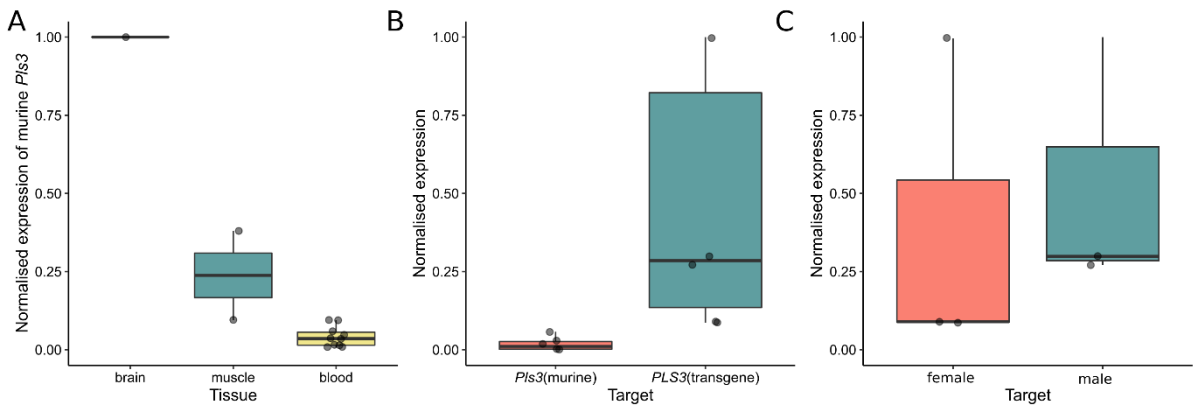


Figure 39: Expression of murine *Pls3* and the human *PLS3* transgene. (A) Murine *Pls3* is expressed in brain tissue (n = 2) of wild-type mice as well as in muscle tissue (n = 2), but not in blood (n = 6). (B) In three female and three male mice, the expression of the murine *Pls3* (red) and the human *PLS3* transgene (green). (C) There was no significant difference in the expression of the human *PLS3* transgene between female and male mice.

We concluded that the *PLS3* transgene was expressed in the transgene mice, while the murine *Pls3* is not expressed in blood of WT and transgene mice.

Next, we analysed the biomarker data. At first, we detected outliers in our data set by applying Tukey's method for outlier detection (Tukey, 1977). By this, we excluded 40 data points. For the analysis of our data, which contained 14 different treatment groups at two time points, we decided to use exclusively non-parametric statistical tests, to neglect the necessity to perform pre-tests to assess normal distribution and equal variances, and more importantly, we were able to use the same test strategy in all comparisons. By this, we probably lost some statistical power, however, the analysis was streamlined and the results more comparable to each other. Using a Kruskal-Wallis-test (KW) we analysed, if two or more groups differed significantly from each other. If this was the case, we conducted *post-hoc* Dunn' test for multiple comparisons to identify which groups differed significantly from each other. If there were only two groups to compare to each other, Wilcoxon Signed-Rank (WSR) tests were used.

5.2.2. No sex-specific differences in SMN and all six plasma biomarkers

Before we began the in-depth analysis of the data sets, we analysed if there were sex-specific differences in the plasma concentration of any of the seven plasma proteins. We compared the concentrations of male and female mice in each of the 14 treatment groups using WSR tests. There were no sex-specific differences in any statistical test. Therefore, we grouped male and female mice in each group.

5.2.2. SMN blood protein concentration

We analysed the concentration of SMN in blood plasma. We split the data set in three groups: untreated animals at P10 (Figure 40A), ASO-treated animals at P10 (Figure 40B), and ASO-treated animals at P21 (Figure 40C). Next, we performed KW-tests to investigate, if two or more groups differed significantly from each other. This was the case in all three data sets. To investigate, which samples differed significantly from each other, we performed *post-hoc* Dunn tests using a Holm correction for multiple comparisons.

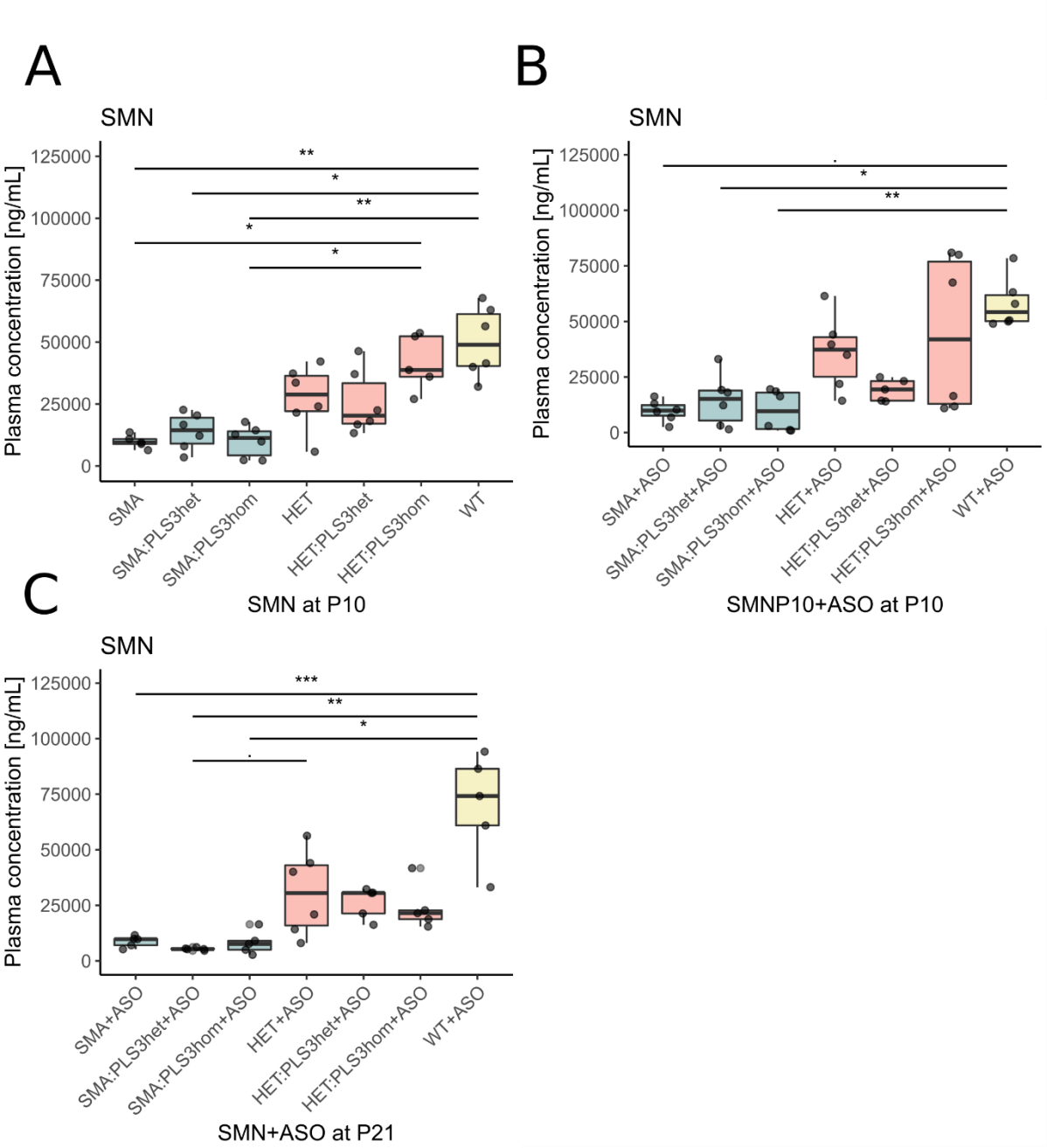


Figure 40: Plasma concentrations of SMN. (A) SMN concentrations in untreated mice at P10, (B) SMN concentrations in ASO-treated mice at P10. (C) SMN concentrations in ASO-treated mice at P21. The variances were analysed by a KW test followed by *post-hoc* Dunn's test for multiple comparisons using a Holm correction. Data are represented as mean ± standard deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

In the untreated animals at P10, we found significant differences between the three SMA groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) and the wild-type animals, indicating that the SMN concentration reflects the disease severity. Furthermore, we found significant differences between SMA and HET:*PLS3*hom animals, as well as between SMA:*PLS3*hom and HET:*PLS3*hom animals. ASO-treated animals at P10 as

well as P21 only showed significant differences between the wild-type and SMA+ASO or SMA:*PLS3*hom+ASO groups. We conclude that the SMN plasma concentration is a useful measure to discriminate severely affected animals (SMA) from the healthy animals at P10 and P21 irrespective of the ASO-treatment.

Furthermore, there were no differences between the three SMA groups (SMA, SMA:*PLS3*het, and SMA:*PLS3*hom) or between the three HET groups (HET, HET:*PLS3*het, and HET:*PLS3*hom). We concluded that the overexpression of the *PLS3* transgene has no impact on the plasma concentration of *SMN*, confirming our previous results (Ackermann et al., 2013). Next, we investigated, if the ASO-treatment influenced the plasma concentration of SMN at P10. We compared the ASO-treated and untreated animals at P10 using the WSR test and found no significant differences. We concluded that the ASO-treatment has no impact on the plasma concentration of *SMN* at P10.

Since we did not detect any impact of *PLS3* overexpression on the SMN level and to increase the power of the statistical tests, we combined the samples of untreated SMA, SMA-*PLS3*het and SMA-*PLS3*hom groups at P10 (pooled SMA, P10) or P21 (pooled SMA, P21), as well as *SMN*-ASO treated SMA, SMA-*PLS3*het and SMA-*PLS3*hom groups at P10 (pooled SMA+ASO, P10). Similarly, we pooled HET and WT animals with or without ASO-treatment at both time points. We performed KW-tests, followed by Dunn tests in the nine pooled groups to investigate if there were significant differences. At both time points, the SMN plasma concentration was always lowest in pooled SMA mice, and highest in the pooled WT mice. In all three comparisons, the concentration in SMA mice differed significantly from the HET mice and from the WT mice, while there was a trend ($P \leq 0.1$) between HET and WT mice. We concluded that the differences in plasma SMN primarily derive from murine *Smn* copies (2 in WT, 1 in HET and 0 in SMA mice) and to a minor extend from the human *SMN2* transgenes (0 in WT, 2 in HET and SMN mice).

Lastly, we investigated how stable the SMN concentration was over time. We compared the SMN concentration in ASO-treated mice at P10 and P21 for each pooled genotype and found no significant differences. We conclude that the concentration of SMN is stable over time and in general a reliable plasma biomarker for SMA (Figure 41).

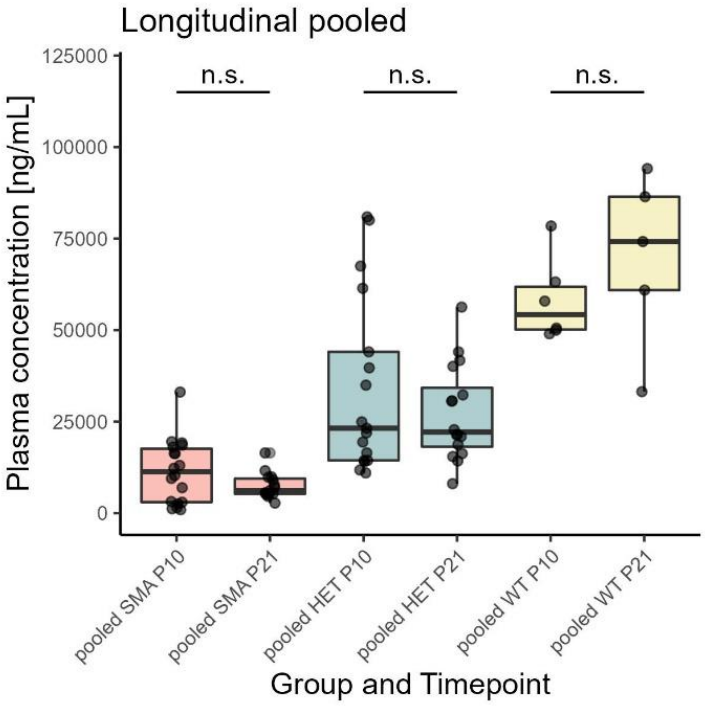


Figure 41: Plasma concentrations of SMN comparing ASO-treated animals at P10 versus P21. The variances were analysed by a KW test followed by a *post-hoc* Dunn's test for multiple comparisons using a Holm correction. Data are represented as mean ± standard deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

Our results demonstrate that SMN levels in the severe SMA mouse model are significantly reduced compared to the wild-type. Furthermore, neither the *SMN*-ASO treatment, nor the overexpression of *PLS3*, did affect the SMN levels in blood of SMA or HET mice, despite increasing the lifespan of the mice in our intermediate mouse model (Hosseinibarkooie et al., 2016). Since the plasma concentrations of SMN were stable over time, we have shown that the SMN is a useful biomarker for SMA in our data set.

5.2.3. Overexpression of PLS3 and ASO treatment have no effect on the concentration of the six biomarkers

Following the same methodology that was applied in the analysis of plasma SMN levels, we continued to analyse the concentrations of the six putative SMA biomarkers (COMP; DPP4, SPP1, CLEC3B, VTN, and AHSG), which have been identified earlier by the BforSMA study (Finkel et al., 2012; Kobayashi et al., 2013).

5.2.4. The PLS3 transgene has no influence on SMN and the six biomarkers

When we compared the plasma concentrations of the six plasma biomarkers in untreated animals at P10, ASO-treated animals at P10 and ASO-treated animals at P21 using KW-tests followed by *post-hoc* Dunn tests, we did generally find no difference between SMA, SMA:*PLS3*het and SMA:*PLS3*hom, except for DPP4. There, in ASO-treated animals at P10, we found a significant difference between SMA:*PLS3*het and SMA:*PLS3*hom mice. Likewise, there was no difference between HET, HET:*PLS3*het and HET:*PLS3*hom animals, respectively. We conclude that the human transgene of *PLS3* does generally not influence the plasma concentrations of the six biomarkers.

5.2.5. Biomarkers for SMA

5.2.5.1. COMP

The general plasma concentration of COMP was severely lower compared to SMN in the range of 100 to 350 ng/mL with decreasing concentrations over time (Figures 42A, B, and C). In the untreated animals at P10, we found in general some differences between the SMA and HET group. In detail, SMA differed from HET:*PLS3*hom and HET:*PLS3*het. SMA:*PLS3*het differed from HET:*PLS3*hom and HET:*PLS3*het. Finally, SMA:*PLS3*het differed from all three HET groups (HET, HET:*PLS3*hom, HET:*PLS3*het). None of the groups was different to the wild-type. In the pooled groups at P10, we found significant differences between untreated and ASO-treated animals.

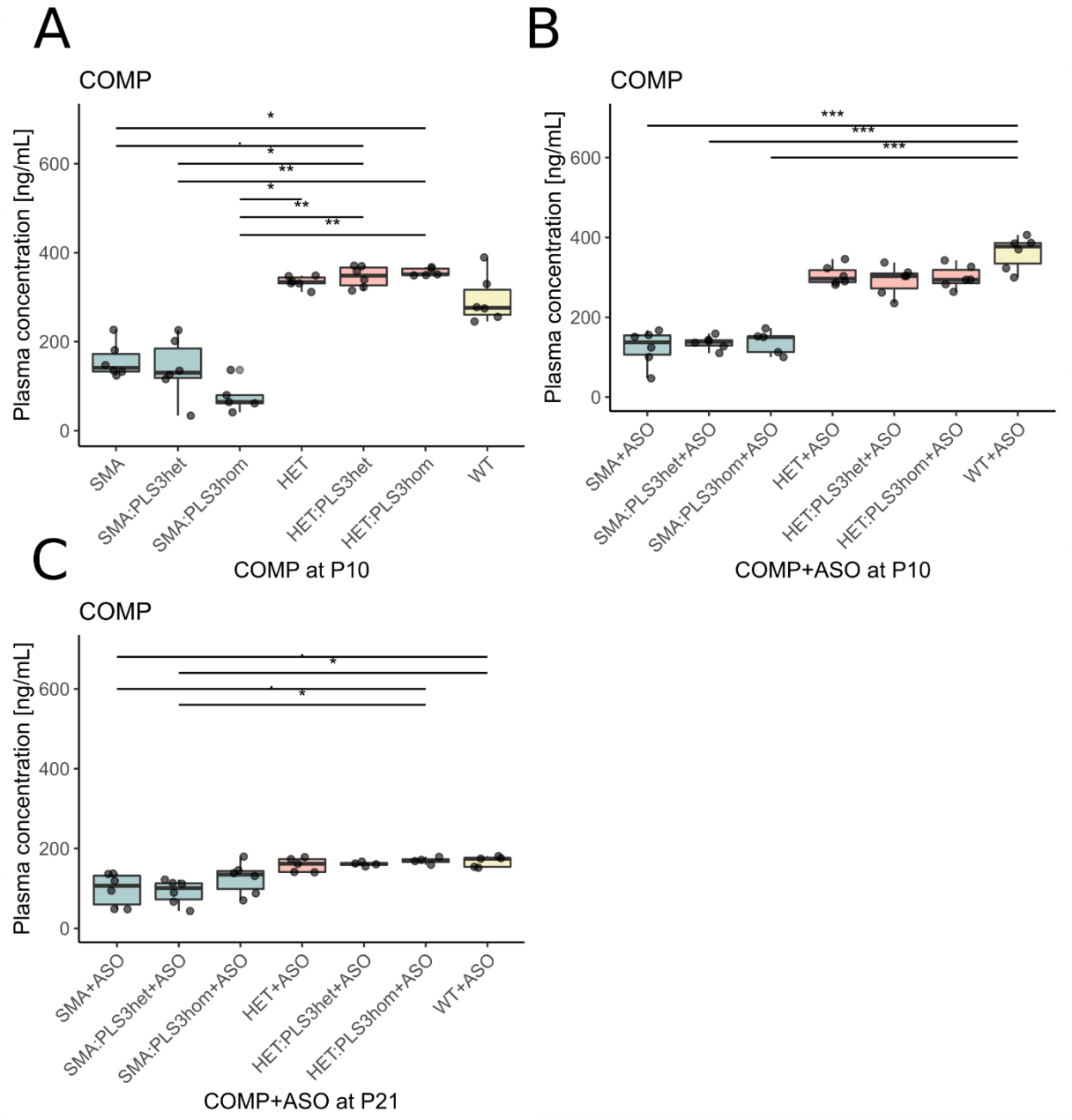


Figure 42: Plasma concentrations of COMP. (A) Concentration of COMP in untreated mice at P10, (B) Concentration of COMP in ASO-treated mice at P10. (C) Concentration of COMP in ASO-treated mice at P21. The variances were analysed by a KW test followed by a *post-hoc* Dunn's test for multiple comparisons using a Holm correction. Data are represented as mean ± standard deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

Furthermore, the biomarker decreased significantly over time in pooled SMA, pooled HET and pooled WT (Figure 43).

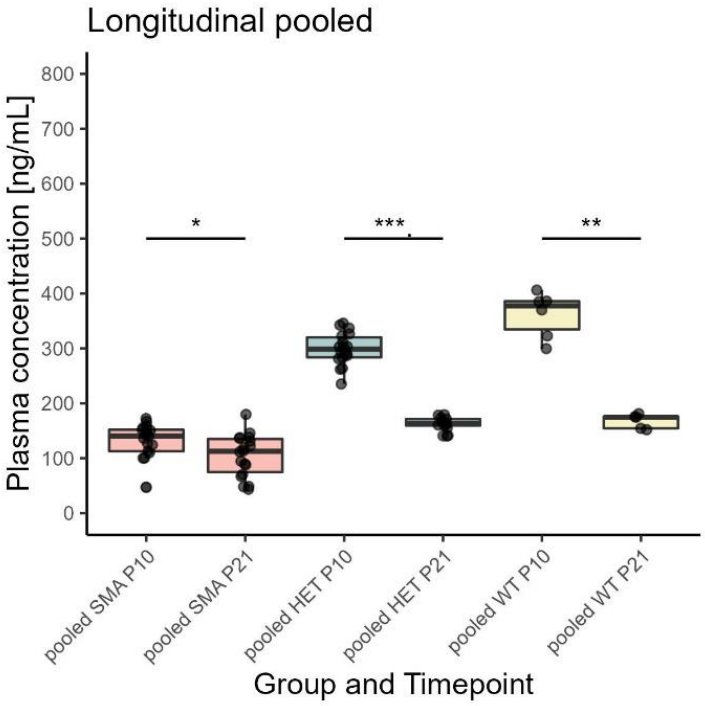


Figure 43: Plasma concentrations of COMP comparing ASO-treated animals at P10 versus P21. The variances were analysed by a KW test followed by a *post-hoc* Dunn's test for multiple comparisons using a Holm correction. Data are represented as mean ± standard deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

As we have shown, the SMN plasma concentration is a reliable indicator of disease severity. Therefore, we compared the COMP plasma levels of each individual mouse to the SMN levels and fitted a linear model into the correlation plots to investigate if there were any linear correlations at P10 or P21, respectively. We found a positive linear correlation at both time points. We found significant linear correlations at P10 ($P \approx 6.16*10^{-10}$, $\rho \approx 0.62$) (Figure 44A), and P21 ($P \approx 2.8*10^{-4}$, $\rho \approx 0.61$) (Figure 44B).

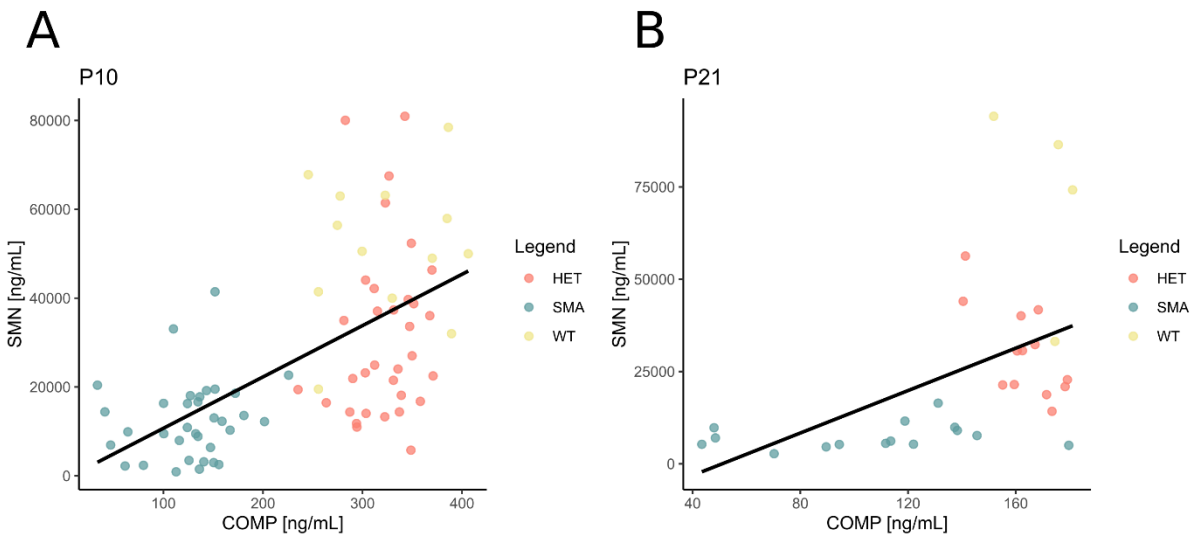


Figure 44: Plasma concentrations of COMP compared to SMN. Linear correlation at (A) P10 and (B) P21. Linear relationships are determined by Spearman's correlation coefficient test deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

Overall, we found that COMP is a well usable plasma biomarker for SMA, as it is associated with the SMA genotype and is correlated with SMN levels at both time points, however, the amount of COMP decreased over time.

5.2.5.2. DPP4

The protein showed very low plasma concentrations in the range of 50 to 150 ng/mL. Untreated animals at P10 showed significant differences between the WT and SMA:*PLS3*het or SMA:*PLS3*hom. Between WT and SMA, there was a trend. This indicates that severely affected animals are characterised by low DPP4 levels, although the effect is less strong than in COMP. Otherwise, there was a significant difference between HET:*PLS3*het and SMA:*PLS3*hom (Figure 45A). This trend is also seen in ASO-treated animals at P10, where SMA:*PLS3*het+ASO animals differed from WT+ASO (Figure 45B). At P21, all three SMA+ASO groups differed significantly from the WT+ASO group (Figure 45C). Therefore, this biomarker shows stronger results at the late time point, because the concentration of the protein increased over time (Figure 46).

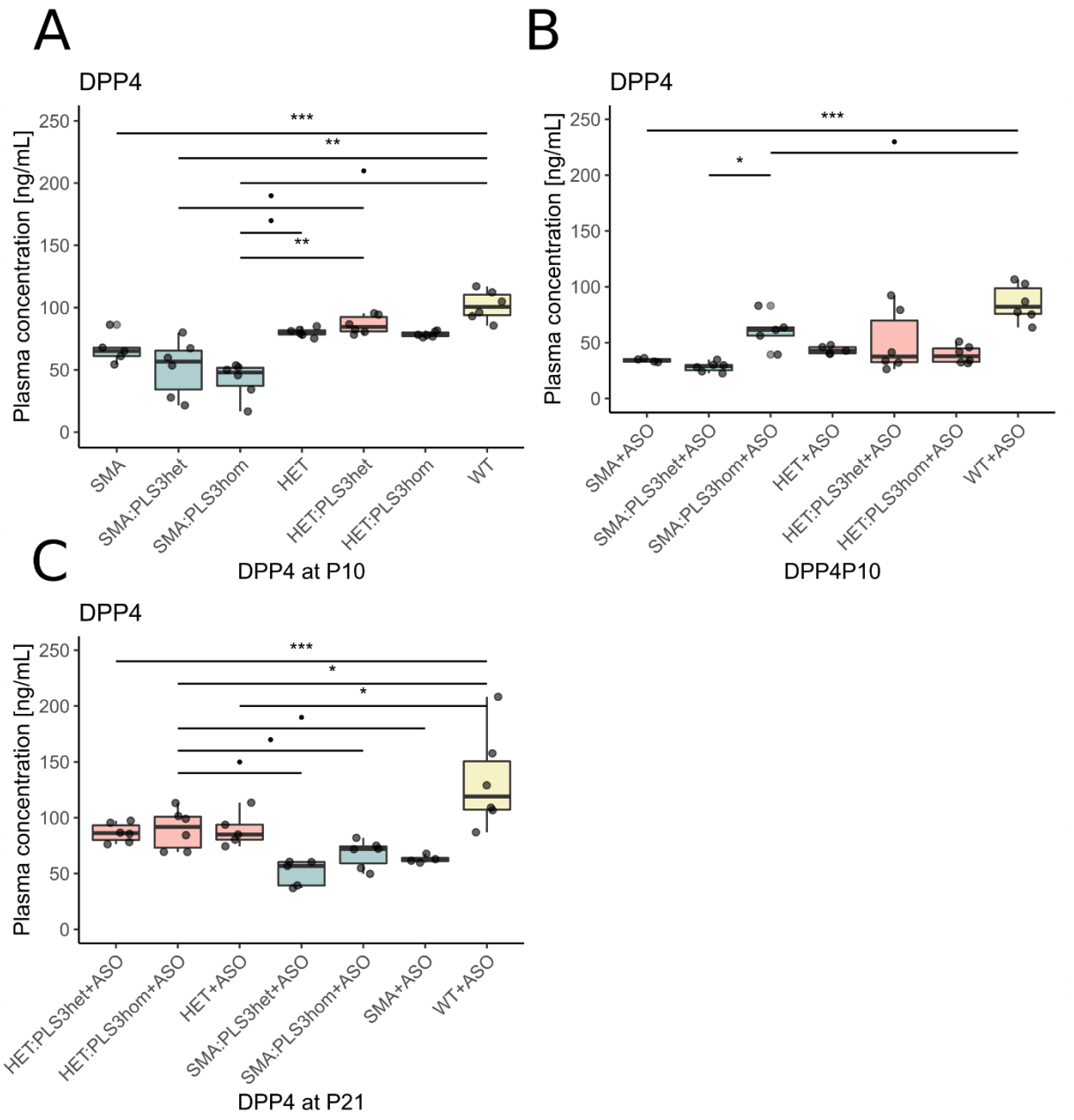


Figure 45: Plasma concentrations of DPP4. (A) Concentration of DPP4 in untreated mice at P10, (B) Concentration of DPP4 in ASO-treated mice at P10, and (C) Concentration of DPP4 in ASO-treated mice at P21. The variances were analysed by a KW test followed by a *post-hoc* Dunn's test for multiple comparisons using a Holm correction. Data are represented as mean ± standard deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

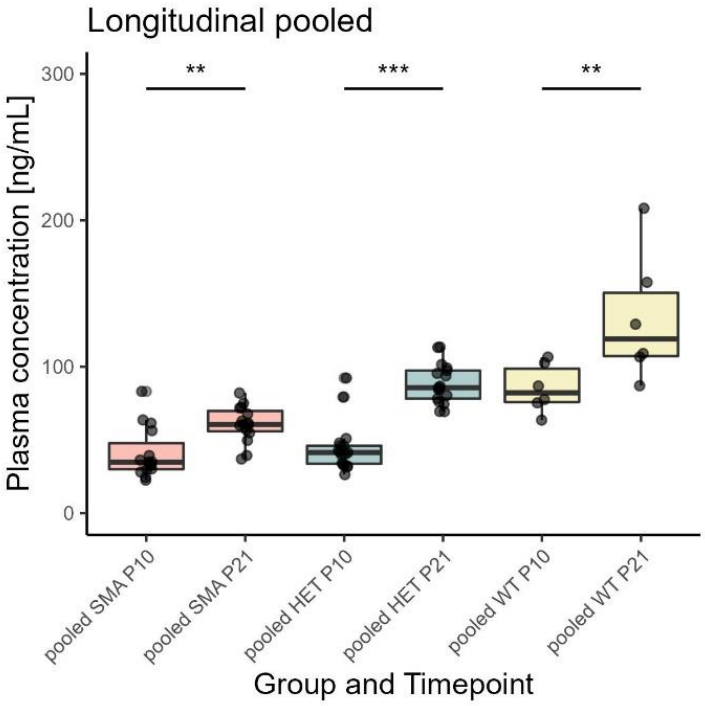


Figure 46: Plasma concentrations of DPP4 comparing ASO-treated animals at P10 versus P21. The variances were analysed by a KW test followed by a *post-hoc* Dunn's test for multiple comparisons using a Holm correction. Data are represented as mean ± standard deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

This is also reflected in the linear correlation plots. We found a significant correlation at both P10 ($P \approx 1.66*10^{-4}$, $\rho \approx 0.41$) (Figure 47A) and P21($P \approx 3.52*10^{-7}$, $\rho \approx 0.82$) (Figure 47B). Overall, we conclude that DPP4 is a useful plasma biomarker especially at later time points.

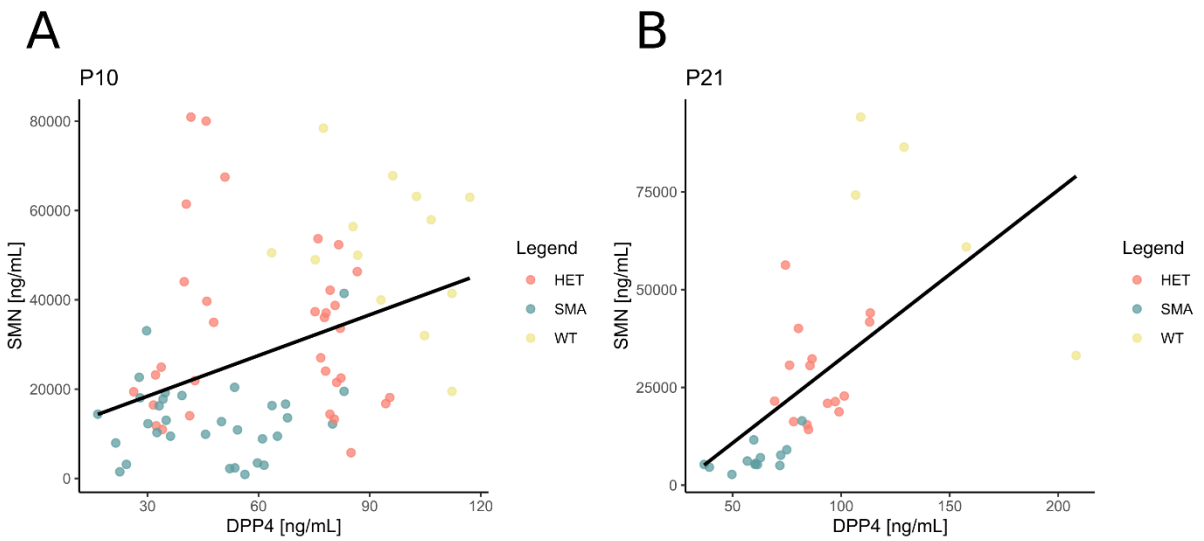


Figure 47: Plasma concentrations of DPP4 compared to SMN. Linear correlation found at (A) P10 and (B) P21. Linear relationships are determined by Spearman's correlation coefficient test deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

5.2.5.3. SPP1

The protein shows a plasma concentration of about 170 to 300 ng/mL. In untreated animals at P10, there was only one significant difference found between SMA:*PLS3*hom and wild-type animals (Figure 48A). In the ASO-treated animals at P10, HET+ASO, SMA:*PLS3*het+ASO and SMA+ASO animals differed from WT+ASO (Figure 48B). The other groups showed in general a huge biological variability. At P21, SMA:*PLS3*hom+ASO animals differed from HET:*PLS3*het+ASO and HET+ASO animals, respectively (Figure 48C). In general, there were no significant differences between ASO-treated and untreated animals.

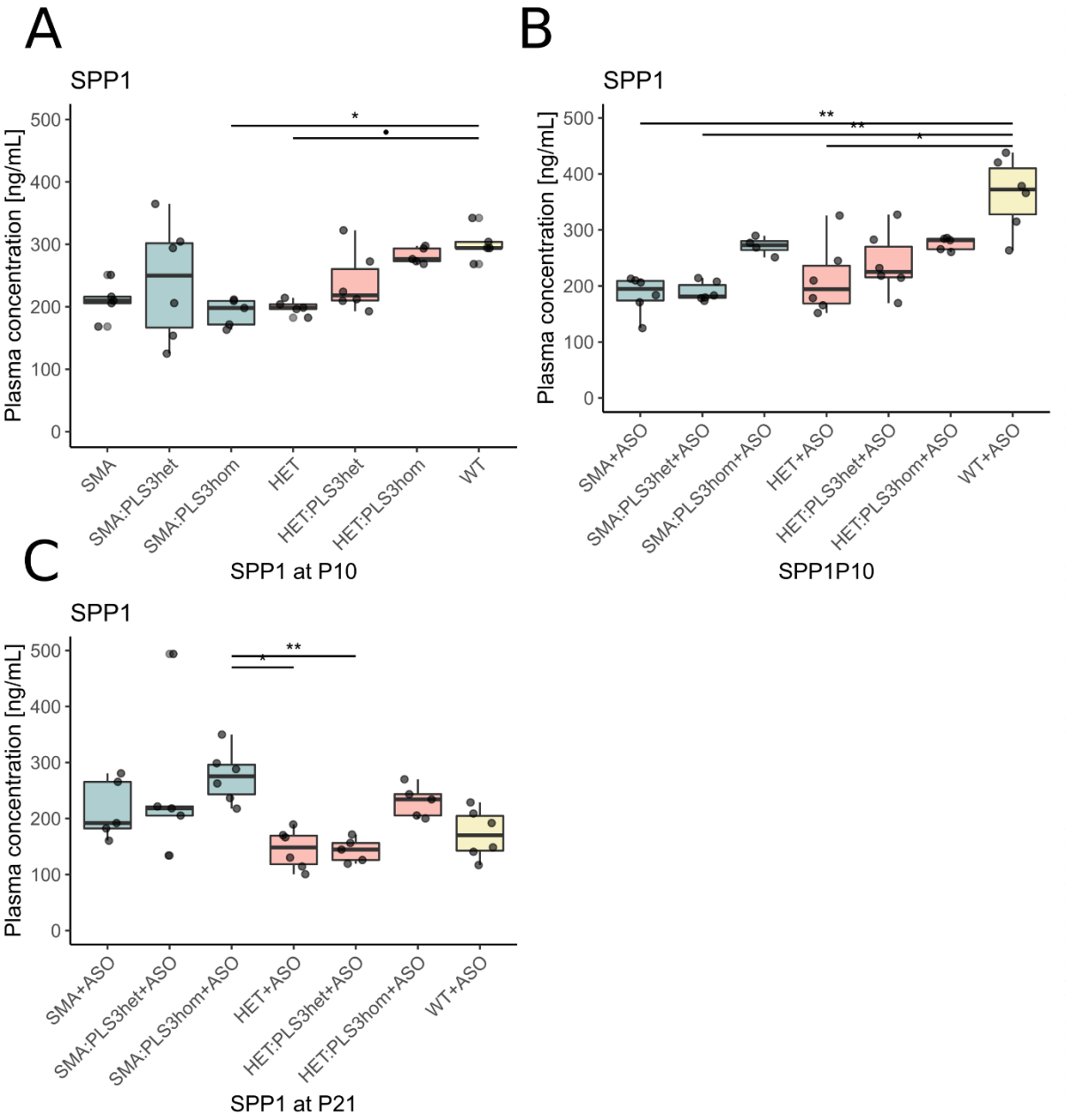


Figure 48: Plasma concentrations of SPP1. (A) Concentration of SPP1 in untreated mice at P10, (B) Concentration of SPP1 in ASO-treated mice at P10. (C) Concentration of SPP1 in ASO-treated mice at P21. The variances were analysed by a KW-test followed by *post-hoc* Dunn's test for multiple comparisons using a Holm correction. Data are represented as mean ± standard deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

The plasma concentration of SPP1 decreased over time in HET and WT animals, but not in SMA animals (Figure 49).

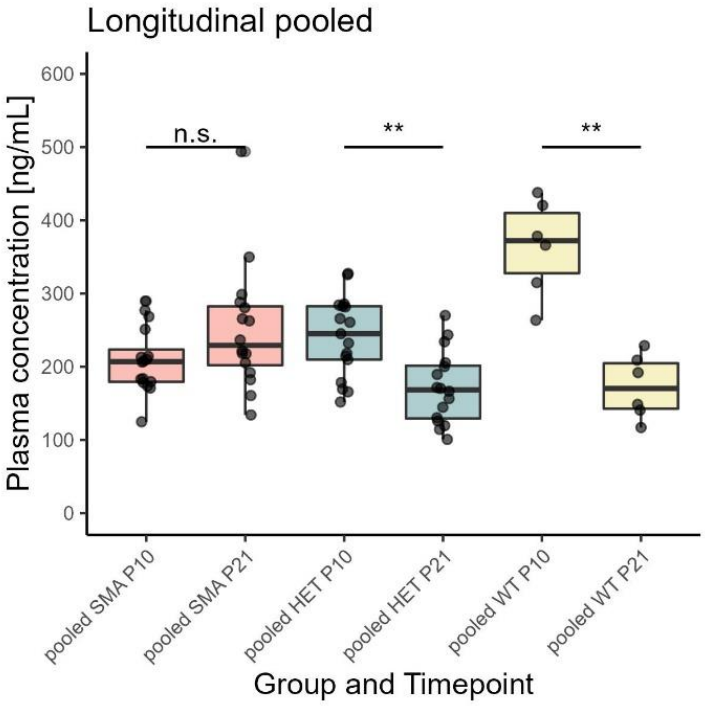


Figure 49: Plasma concentrations of SPP1 comparing ASO-treated animals at P10 versus P21. The variances were analysed by a KW-test followed by *post-hoc* Dunn's test for multiple comparisons using a Holm correction. Data are represented as mean ± standard deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

This is also reflected in the linear comparisons to SMN. We found intermediate correlations at P10 ($P \approx 7.45*10\text{-}6$, $\rho \approx 0.49$) (Figure 50A) and a negative correlation at P21 ($P \approx 0.014$, $\rho \approx -0.42$) (Figure 50B). This renders SPP1 as insufficient plasma biomarker in our data set.

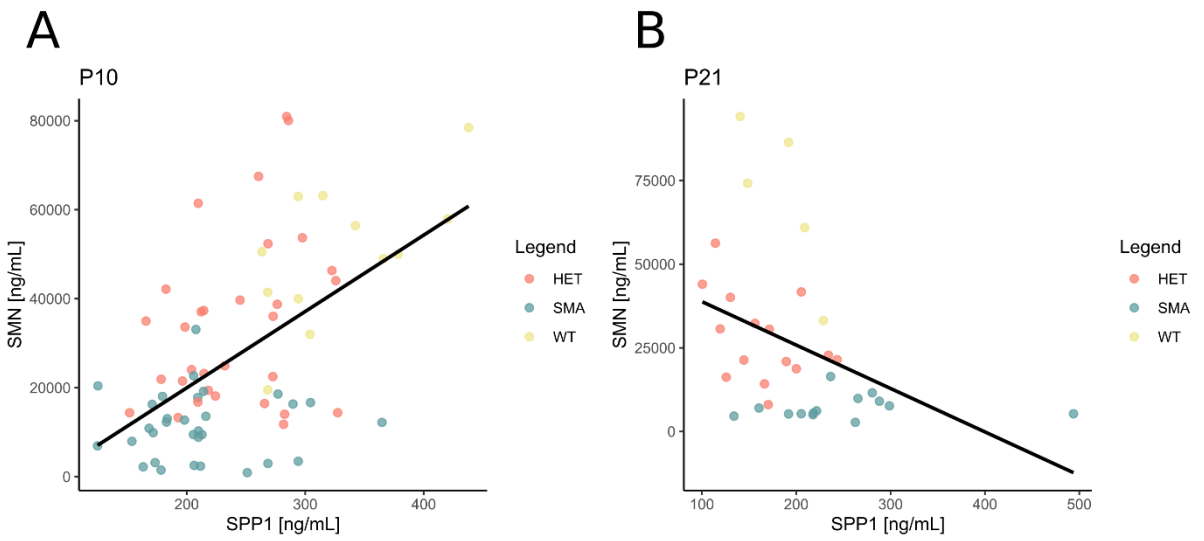


Figure 50: Plasma concentrations of SPP1 compared to SMN. Linear correlation found at (A) P10 and (B) P21. Linear relationships are determined by Spearman's correlation coefficient test deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

5.2.5.4. CLEC3B

This plasma biomarker showed the lowest plasma concentrations in the range of 10 to 20 ng/mL. Furthermore, the concentrations were influenced by the ASO-treatment in all pooled groups. At P10 in untreated animals, there was only one difference between SMA:*PLS3*hom and HET:*PLS3*het animals (Figure 51A). The ASO-treated SMA animals differed significantly from the WT+ASO group at P10. Furthermore, there were differences between SMA:*PLS3*hom+ASO and HET+ASO or HET:*PLS3*hom+ASO, respectively (Figure 51B). At P21, the plasma concentrations between the genotypes were not differentiable, except for HET:*PLS3*hom+ASO (Figure 45C).

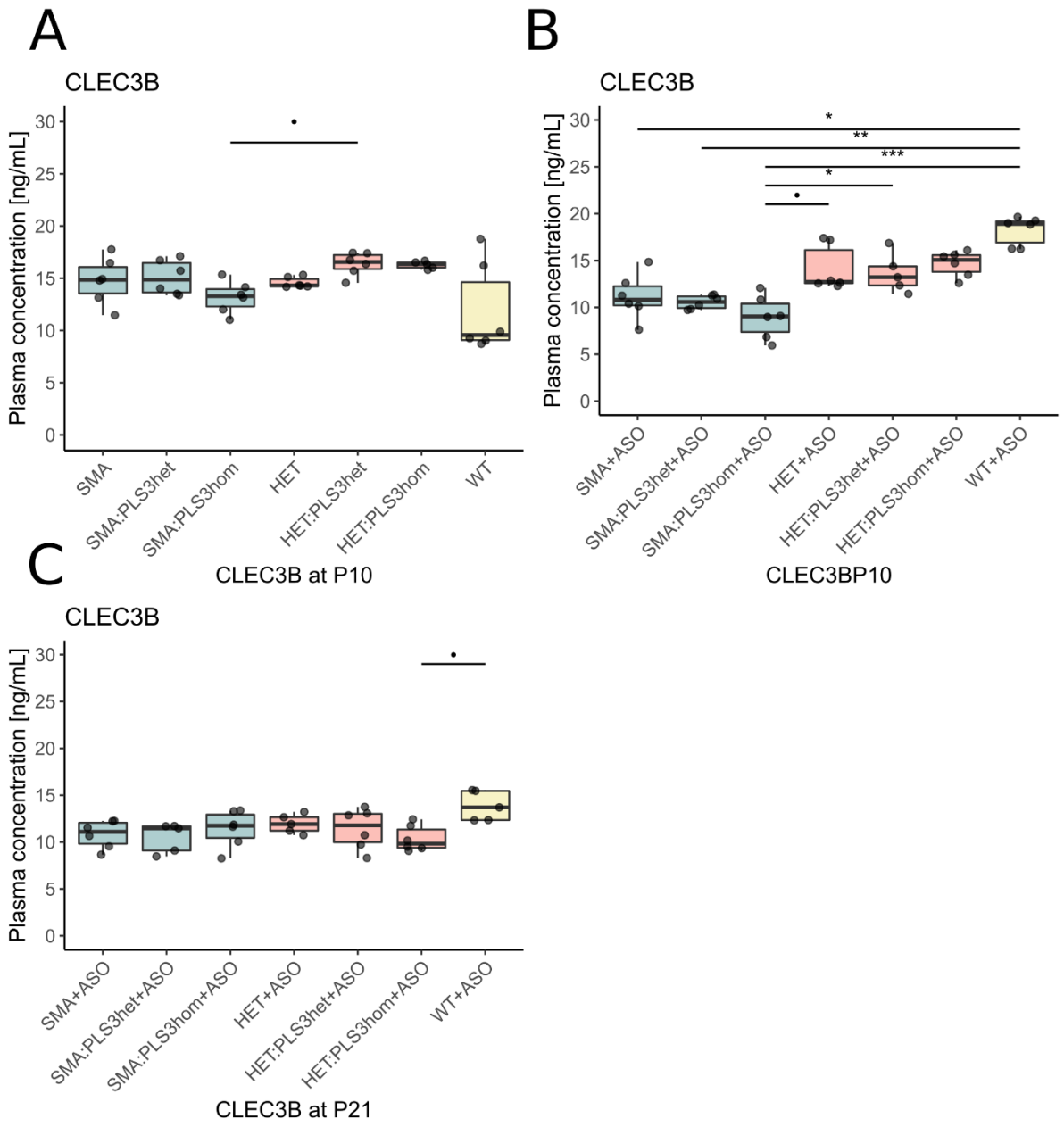


Figure 51: Plasma concentrations of CLEC3B. (A) Concentration of CLEC3B in untreated mice at P10, (B) Concentration of CLEC3B in ASO-treated mice at P10. (C) Concentration of CLEC3B in ASO-treated mice at P21. The variances were analysed by a KW-test followed by *post-hoc* Dunn's test for multiple comparisons using a Holm correction. Data are represented as mean ± standard deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

The longitudinal comparison shows that the protein plasma concentration is significantly decreased at P21 compared to P10 in HET and WT animals (Figure 52).

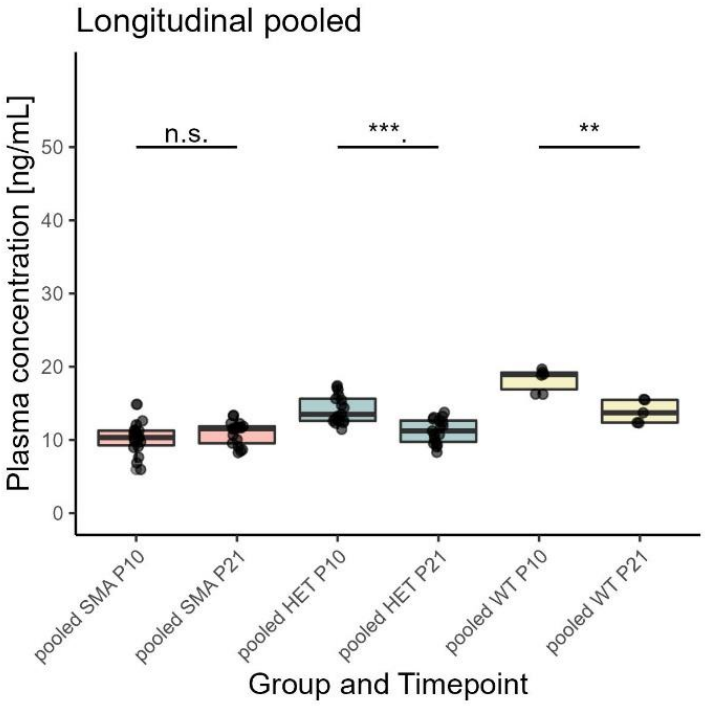


Figure 52: Plasma concentrations of CLEC3B comparing ASO-treated animals at P10 versus P21. The variances were analysed by a KW-test followed by *post-hoc* Dunn's test for multiple comparisons using a Holm correction. Data are represented as mean ± standard deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

Furthermore, the linear comparison with SMN showed only moderate correlation at P10 ($P \approx 4.3*10^{-4}$, $\rho \approx 0.37$) (Figure 53A) and P21 ($P \approx 2.05*10^{-2}$, $\rho \approx -0.4$) (Figure 53B). We therefore conclude from our data, that CLEC3B is not a reliable biomarker for SMA.

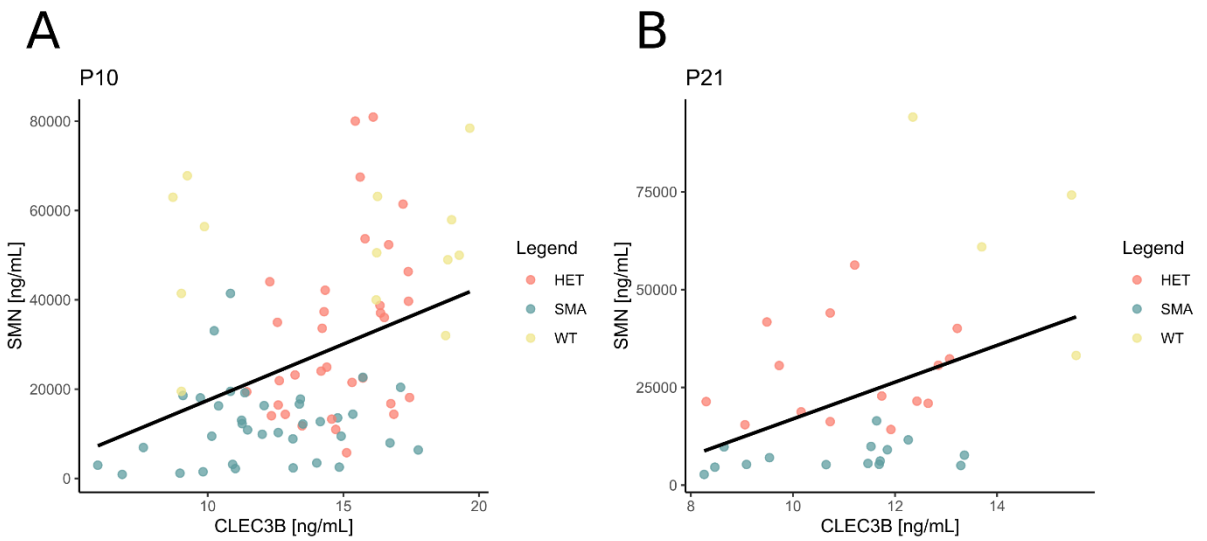


Figure 53: Plasma concentrations of CLEC3B compared to SMN. Linear correlation at (A) P10 and (B) P21. Linear relationships are determined by Spearman's correlation coefficient test deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

5.2.5.5. VTN

This plasma protein showed relatively strong plasma concentrations in the range of 700 to 2700 ng/mL. In untreated animals at P10, we found that the three HET groups had stronger concentrations than the wild-type animals. The HET:*PLS3*het showed the strongest VTN plasma levels and differed significantly from SMA, WT, and SMA:*PLS3*hom animals. Furthermore, HET:*PLS3*hom animals differed from SMA:*PLS3*hom animals (Figure 54A). The ASO-treated HET animals at P10 showed plasma levels similar to the WT animals. There were significant differences between SMA:*PLS3*hom+ASO and both HET:*PLS3*het+ASO and HET:*PLS3*hom+ASO. The SMA:*PLS3*het+ASO animals differed from the WT+ASO animals (Figure 54B). At P21, the data did not pass the KW-test and no differences between the groups was obtained (Figure 54C). The strong variation in HET animals is also shown in the comparison of untreated and ASO-treated animals at P10, where a significant difference is found.

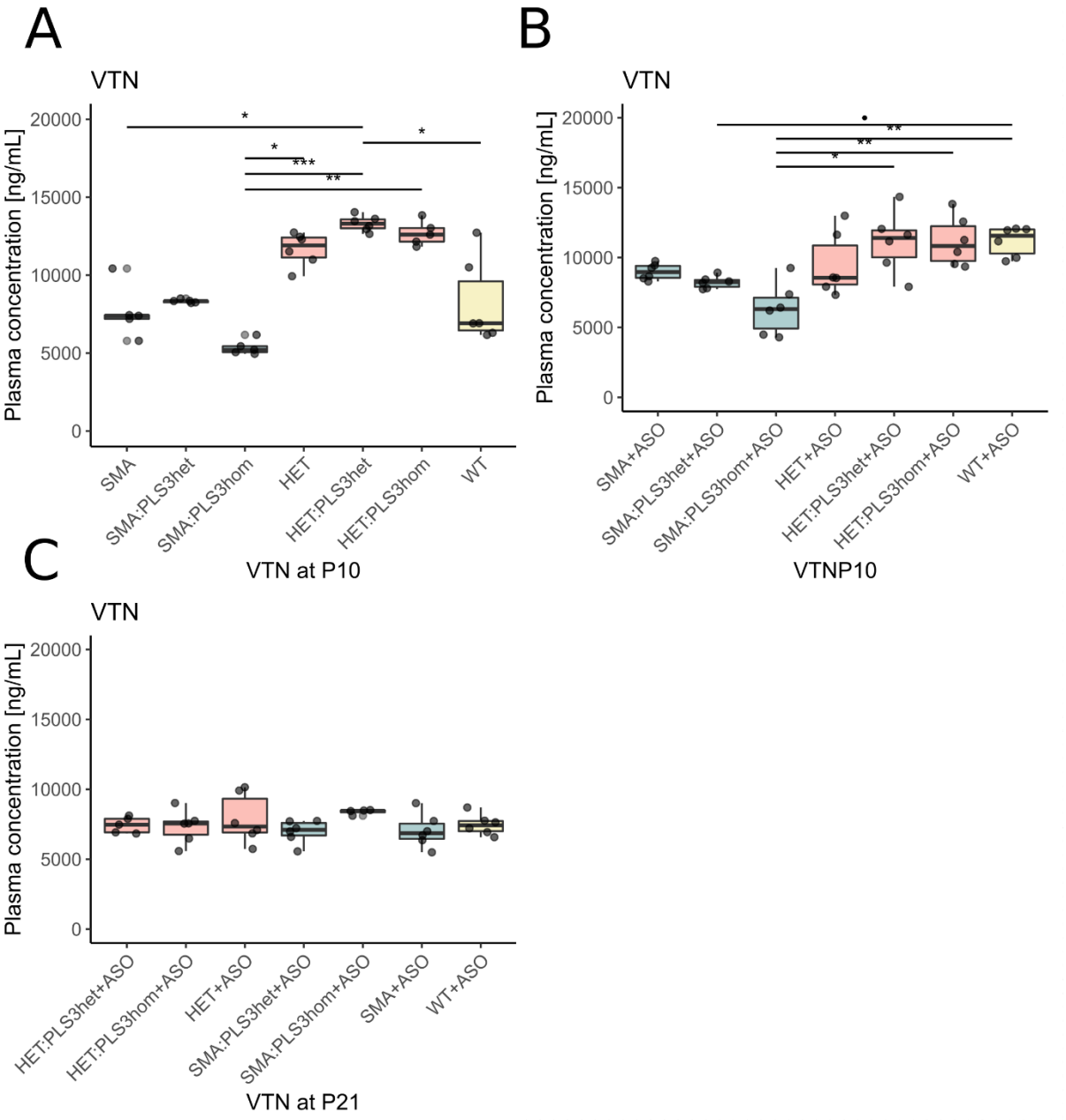


Figure 54: Plasma concentrations of VTN. (A) Concentration of VTN in untreated mice at P10. (B) Concentration of VTN in ASO-treated mice at P10. (C) Concentration of VTN in ASO-treated mice at P21. The variances were analysed by a KW-test followed by *post-hoc* Dunn's test for multiple comparisons using a Holm correction. Data are represented as mean ± standard deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

Furthermore, the plasma concentration of VTN decreased over time in HET and WT animals (Figure 55).

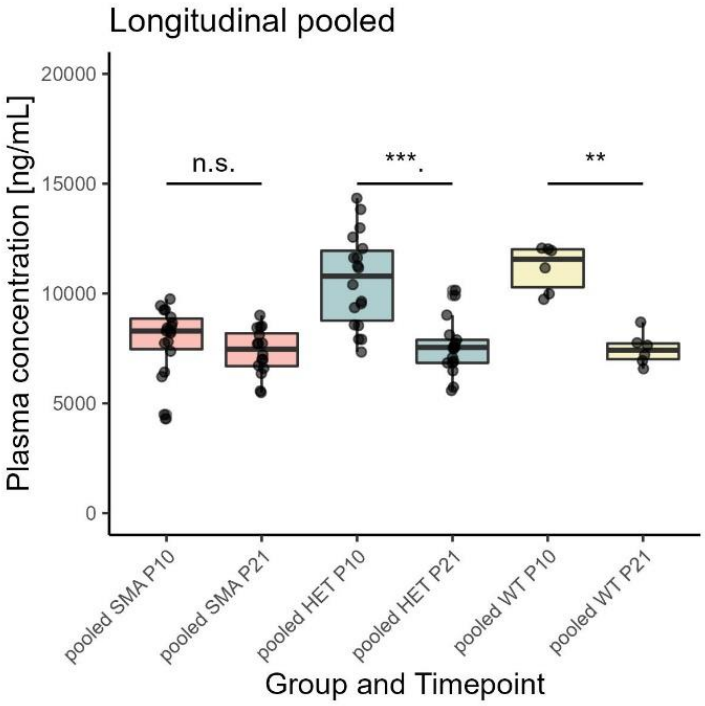


Figure 55: Plasma concentrations of VTN comparing ASO-treated animals at P10 versus P21. The variances were analysed by a KW-test followed by *post-hoc* Dunn's test for multiple comparisons using a Holm correction. Data are represented as mean ± standard deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

The linear correlation with SMN was intermediate at P10 ($P \approx 1.36*10^{-4}$, $\rho \approx 0.41$) (Figure 56A) and not significant at P21 ($P \approx 0.32$, $\rho \approx 0.17$) (Figure 56B). Our data does not support VTN as reliable SMA biomarker in this mouse model.

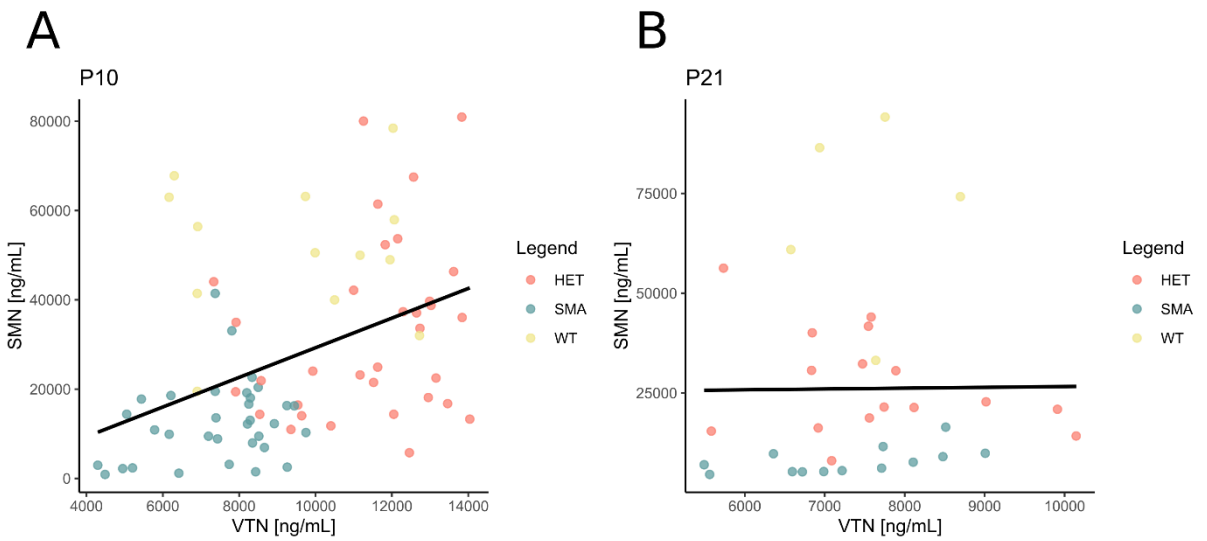


Figure 56: Plasma concentrations of VTN compared to SMN. Linear correlation at (A) P10 and (B) P21. Linear relationships are determined by Spearman's correlation coefficient test deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

5.2.5.6. AHSG

Finally, AHSG showed intermediate expression in blood plasma with ranges between 80 to 250 ng/mL. The plasma concentration was affected by the ASO-treatment in SMA animals. Interestingly, the plasma concentrations were strongest in SMA animals and lowest in HET animals. There were significant differences in untreated animals at P10 between SMA and HET or HET:*PLS3*hom animals, but not to the wild-type (Figure 57A). In ASO-treated animals at P10, SMA+ASO animals differed from HET:*PLS3*het+ASO and HET:*PLS3*hom+ASO. SMA:*PLS3*het+ASO animals differed from both HET:*PLS3*het+ASO and HET:*PLS3*hom+ASO animals (Figure 57B). At P21, there was only one significant difference between HET+ASO and SMA:*PLS3*hom+ASO (Figure 57C).

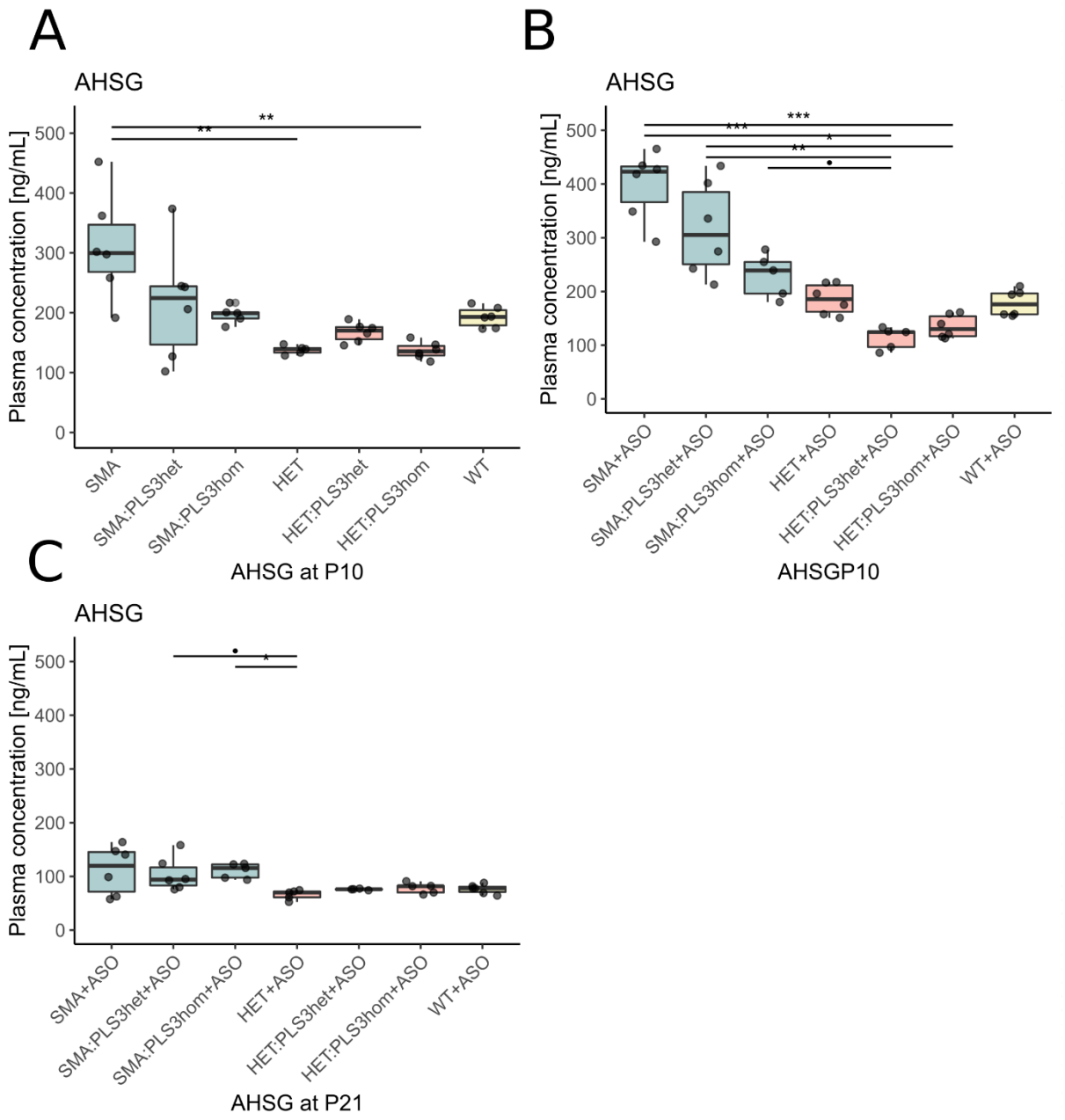


Figure 57: Plasma concentrations of AHSG. (A) Concentration of AHSG in untreated mice at P10. (B) Concentration of AHSG in ASO-treated mice at P10. (C) Concentration of AHSG in ASO-treated mice at P21. The variances were analysed by a KW-test followed by *post-hoc* Dunn's test for multiple comparisons using a Holm correction. Data are represented as mean ± standard deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

Furthermore, the plasma concentration of AHSG is instable over time in all pooled groups (Figure 58).

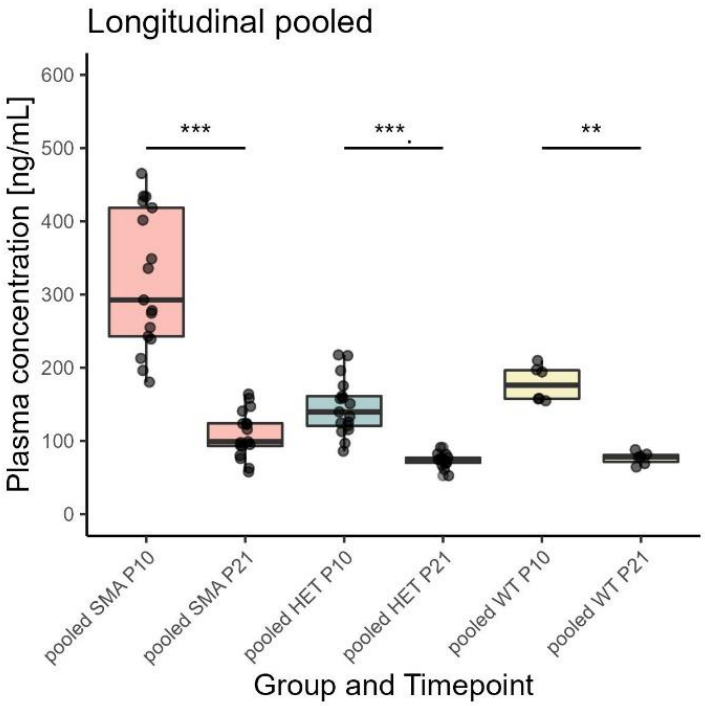


Figure 58: Plasma concentrations of AHSG comparing ASO-treated animals at P10 versus P21. The variances were analysed by a KW-test followed by post-hoc Dunn's test for multiple comparisons using a Holm correction. Data are represented as mean ± standard deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

At both time points, there was a negative moderate correlation with SMN, which was significant at P10 ($P \approx 0.007$, $\rho \approx -0.37$) (Figure 59A) and insignificant at P21 ($P \approx 0.058$, $\rho \approx -0.32$) (Figure 59B). As the protein is involved in bone development, the upregulation might be due to delayed development in severely affected SMA mice (Brylka et al., 2017). Our data suggest that AHSG is a moderately useful SMA biomarker.

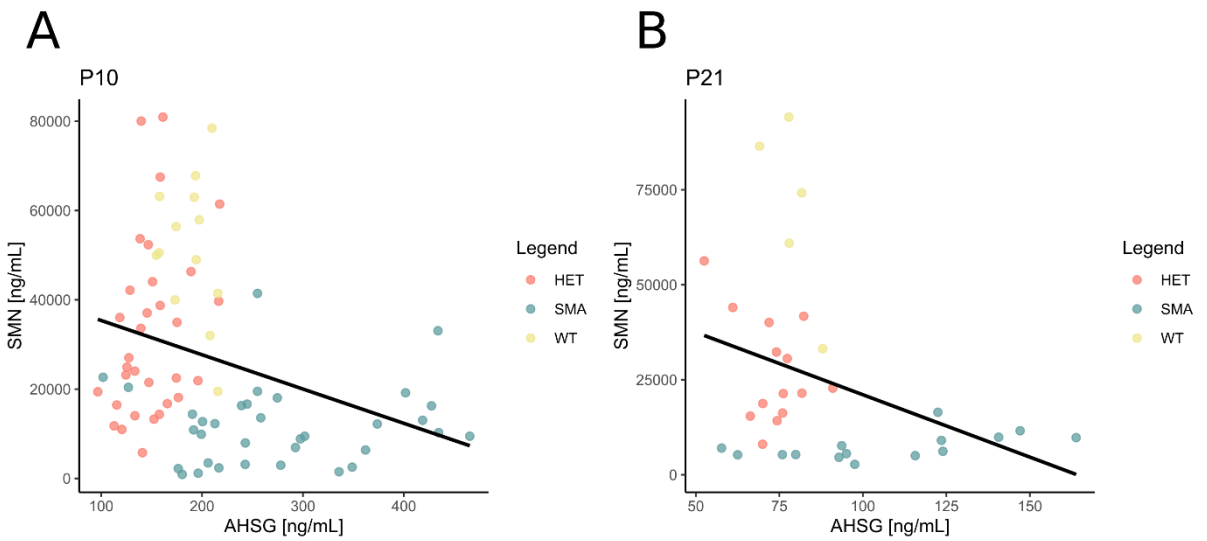


Figure 59: Plasma concentrations of AHSG compared to SMN. Linear correlation at (A) P10 and (B) P21. Linear relationships are determined by Spearman's correlation coefficient test deviation (*P < 0.05; **P < 0.01; ***P < 0.001). SMA mice groups (SMA, SMA:*PLS3*het, SMA:*PLS3*hom) are shown in green; HET mice groups (HET, HET:*PLS3*het, HET:*PLS3*hom) are shown in red; wild-type control mice are shown in yellow.

In summary, pooled SMA groups showed highly significant decreased levels for three biomarkers, COMP, DPP4 and SPP1, at both time points. In contrast, CLEC3B, VTN and AHSG do not appear as reliable biomarkers in this mouse model based on our data. Overexpression of *PLS3* by a human transgene had no influence on the biomarkers, while low-dose systemically administered *SMN*-ASO treatment influenced the plasma levels in some cases (Table 8).

Table 8: Results of the biomarker study (modified after Strathmann et al., 2018).

	SMN	COMP	DPP4	SPP1	CLEC3B	VTN	AHSG
Plasma levels	high				very low	inter-mediate	low
Sex-specific differences	no	no	no	no	no	no	no
Responsiveness to SMN ASO	no	yes	yes	no	in SMA	in HET	yes
Responsiveness to *PLS3* transgene	no	no	at P10	no	no	no	no
Correlation with SMA at P10	yes	yes	yes	yes	yes	yes	anti
Correlation with SMA at P21	yes	yes	yes	anti	yes	no	no
Responsiveness to SMA	very good	good	good	moderate			
Stable over time	yes	no	no	no	no	no	no

5.2.5.7. Correlations of protein levels between the six biomarkers

To determine if the biomarkers belong to discrete mechanistic pathways or if they work independently of each other, we used the previously introduced setup to correlate all six putative biomarkers with each other (Table 9, 10).

Table 9: Correlation of all six biomarkers. The table gives the Spearman correlation coefficient ρ for each pair of biomarkers at P10 and if the correlation is significant (•P>0.1; *P < 0.05; **P < 0.01; ***P < 0.001).

P10	COMP	SPP1	CLEC3B	VTN	AHSG	DPP4
COMP		0.43 ***	0.56 ***	0.66 ***	-0.59 ***	0.19 •
SPP1			0.03 n.s.	0.05 n.s.	-0.11 n.s.	0.48 n.s.
CLEC3B				0.78 ***	-0.48 ***	-0.27 **
VTN					-0.44 ***	-0.29 **
AHSG						0.13 n.s.
DPP4						

Table 10: Correlation of all six biomarkers. The table gives the Spearman correlation coefficient ρ for each pair of biomarkers at P21 and if the correlation is significant (•P>0.1; *P < 0.05; **P < 0.01; ***P < 0.001).

P21	COMP	SPP1	CLEC3B	VTN	AHSG	DPP4
COMP		-0.16 n.s.	0.45 *	0.55 **	-0.13 n.s.	0.80 ***
SPP1			0.10 n.s.	0.30 n.s.	0.65 ***	-0.33 •
CLEC3B				0.34 •	0.23 n.s.	0.24 n.s.
VTN					0.42 *	0.46 *
AHSG						-0.26 n.s.
DPP4						

At P10, the strongest correlation is found between COMP, VTN and CLEC3B. Furthermore, good correlations (ρ ≥ 0.5) were found between COMP and DPP4 as well as between SPP1 and DPP4. AHSG shows at P10 negative Spearman correlation coefficients. It shows weak negative correlations with COMP, VTN and CLEC3B. Interestingly, it shows the strongest positive correlation with SPP1 at P21. Also, a good correlation (ρ ≥ 0.5) was found between COMP and DPP4 at P21. These data suggest i) that COMP, CLEC3B and VTN may share similar mechanistic pathways, ii) DPP4, SPP1 are only weakly correlated with the other biomarkers and may work independently of each other and iii) AHSG is very weakly anti-correlated with COMP, CLEC3B and VTN and completely independent of DPP4 and SPP1 at P10.

6. Discussion

6.1. Part I: Expression of PLS3 in health and disease

Here we have shown that *PLS3* is able to escape XCI in iPSC-derived spinal MNs. The expression of *PLS3* correlates with the copy number of the *DXZ4* macrosatellite in females, but not in males. *PLS3* high expressing females carry at least one *DXZ4* allele of increased copy number (>70 repeats). However, there are rare cases of males, which show an increased expression of *PLS3*. The transcriptional regulator CHD4 is able to bind to the promoter of *PLS3* and induces expression of the gene in males and females.

6.1.1. PLS3 escapes XCI in spinal MNs

We have identified escape genes in spinal MNs (including *PLS3*) by calling of biallelic, exonic SNVs in transcriptome data. We identified two biallelic SNVs in the exons 11 and 12 of *PLS3* with a SNV ratio of about 0.4 in both SMA-discordant families. The differential expression analysis showed a doubling of the *PLS3* expression in asymptomatic females compared to their SMA-affected brothers. With our approach, we have shown that *PLS3* is able to escape from XCI in spinal MNs and that both alleles are transcribed in similar proportions. This is in line with the previous observation that *PLS3* is a cell-type-specific facultative escape X-linked gene (Balaton et al., 2015; Carrel and Willard, 2005; Cotton et al., 2013; Yang et al., 2010). A comparison of our list of escape genes has shown that there is a strong discordance of escape genes that have been identified by different studies (Figure 11). Only 2.3% of escape genes have been found by all mentioned studies. This can be explained by the fact that the escape from XCI is highly tissue specific (Berletch et al., 2015). The escape genes were identified in several different tissues (motor neurons, fibroblasts, B lymphocytes, and leucocytes) and by different methods. Interestingly, the number of genes that were identified in each tissue are very similar, ranging between 75 and 114 genes. This indicates that in each tissue a similar ammount of genes escape XCI. However, our approach of escape gene identification has some limitations. First of all, we were only able to detect genes that contain biallelic SNVs that must reside in exons. Homozygous SNVs are not informative, because they do not allow to discriminate the two alleles. In addition to that, escape genes that are not actively transcribed at the time point of RNA extraction cannot be identified, although they might have been

expressed at an earlier time point. These sources of error render the list of identified escape genes incomplete and represent only one specific time point. Another source of error may be pseudogenes or copy genes, which may align to the genetic locus of target genes. A scenario where those misaligned reads contain genetic variants would facilitate the false-positive identification of escape genes.

To circumvent these limitations, the two SNVs that were identified in *PLS3* were validated by multiple methods. We identified the SNVs in transcriptomes of the spinal MNs, FBs and EBV cells, and confirmed them by Sanger sequencing. The differential expression analysis was performed by a pseudoalignment method called KALLISTO, while the transcriptomes were aligned by HISAT2 (see methods 7.2.2.4) (Bray et al., 2016; Kim et al., 2019). We manually examined the bam files of the MN transcriptomes and checked the MAPQ-scores. This score quantifies the probability that a read is aligned to the wrong genetic locus (Li et al., 2008). In HISAT2, the maximum mapping quality is 60, which means that the particular read is mapped one time to the reference genome (Kim et al., 2019). An examination of the bam files has shown that the reads, which contain the two *PLS3* SNVs were mapped uniquely to the reference genome (MAPQ-score = 60). However, we screened all 48 available transcriptome data sets (table 1) for exonic biallelic *PLS3* SNVs. The two SNVs have not been found in any other sample than families 1 and 2, nor did we identify other informative genetic variants. Therefore, it is not possible to repeat and validate our results with spinal MNs derived from FBs of our study participants.

6.1.2. PLS3 expression correlates with DXZ4 copy number

PLS3 and *DXZ4* are located in close proximity on Xq23 only 500 kb apart from each other next to the boundary of the two megadomains on the X_i (Rao et al., 2014). We found a significant correlation between the copy number of the macrosatellite and the expression of *PLS3* in females (Figures 12B, 16, and 19). *DXZ4* consists of hypomethylatedmethylated open chromatin marks on the inactivated X_i (Chadwick, 2008; Giacalone et al., 1992). This chromosomal conformation could explain the correlation between the *DXZ4* copy number and the expression levels of *PLS3*, as the sheer size of the macrosatellite would influence the localisation of the genomic locus in the nucleus. Thereby, the copy number of the macrosatellite would affect the expression of neighbouring genes, such as *PLS3*. If this is indeed the case, it would mean that the escape of *PLS3* is favoured, if the X_i harbours a *DXZ4* allele with a high

copy number. Our data set supports this idea, as we found at least one *DXZ4* allele of increased copy number (> 70 repeats) in EBV cells with high expression of *PLS3*. Furthermore, there is a significant difference in the absolute *DXZ4* copy number between *PLS3* high and low expressers. Additionally, samples with large average *DXZ4* copy numbers seem to be more prone to changes in the expression of *PLS3* (Figure 23). Based on these results and to gain more insight, we analysed the XI-status of nine cell lines using the methylation-sensitive X-inactivation assays. However, a clear difference in the XI-status was found only in three pairs of cell lines. (Figure 25). One explanation might be that the measured differences in the XI-status were not strong enough to change the expression of *PLS3* significantly. Another explanation might be that the escape of *PLS3* from XCI is primarily influenced by the copy number of *DXZ4* and independent of the general XI-status of the X-chromosomes. This idea is supported by a recently published bioinformatics tool, DeepLoop, which is able to enhance chromatin interaction mapping by applying bias correction and deep-learning-based signal enhancement. This tool was used to analyse a data set from GM12878 cells, in which the paternal X-chromosome is inactivated. They identified escape genes near the *DXZ4* locus on the X_i and conclude that those escape regions are mechanistically coupled to the formation of the megadomains (Zhang et al., 2022). One important limitation of our study might be that the selection of cells is biased by previous knowledge about the expression of *PLS3* in our sample library. The *DXZ4* copy numbers as well as the expression of *PLS3* were measured in samples that were either identified to be asymptomatic individuals from SMA-discordant families, SMA-affected individuals with various *PLS3* expression levels as well as healthy individuals from our EBV cell library. Therefore, our data sets represent a study population with a strong bias towards SMA, while healthy individuals are underrepresented. However, increased expression of *PLS3* in blood is a rare event and our previous data show that 95% of healthy individuals do not express the gene (Oprea et al., 2008). Interestingly, the *DXZ4* copy numbers measured by Molecular Combing show strong similarities to previously published data. Other research groups found copy numbers between 12 and 100 repeats, measured by extended DNA fibre FISH, while we found copy numbers between 27 and 110 repeats (Giacalone et al., 1992; Tremblay et al., 2011).

6.1.3. CHD4 is an epigenetic transcriptional regulator of PLS3

One pitfall of our analysis was that the *PLS3* expression seems to be modified by several mechanisms at once, such as escape from XCI, and changes in the XI-status. Those layers of regulation primarily influence the expression of *PLS3* in female cell lines. The fact that we identified rare cases of *PLS3* overexpression in lymphoblastoid cell lines of males was enigmatic and pointed towards additional layers of *PLS3* regulation. Furthermore, we found a general, but lower expression of *PLS3* in spinal MNs of SMA-affected siblings of asymptomatic women (Figure 7C, D). By analysis of transcriptome data from male EBV cells followed by knock-down, overexpression experiments, as well as ChIP-qPCR and luciferase-promoter assays, we identified CHD4 as transcriptional regulator of *PLS3* in EBV and HEK293T cells. To reduce the number of parameters and to examine the influence of CHD4 without the effect of changes of the XI-status, we performed knock-down experiments in EBV cells of both sexes. We assumed that the XI-status would not change between *CHD4* and mock siRNA transfected cells. The XI-status would only influence the *PLS3* expression in females, as males are hemizygous for *PLS3* and *DXZ4*. siRNA-mediated knock-down of *CHD4* in EBV cells confirmed our finding that CHD4 is an epigenetic regulator of *PLS3*. Similar to *PLS3*, high *CHD4* levels have been found in colorectal, breast, pancreatic, and lung cancer (Chang et al., 2021; Wang et al., 2020; Xin et al., 2020; Xu et al., 2020). Interestingly, there is no difference in the expression of *CHD4* in MNs between asymptomatic and SMA-affected siblings in the two SMA-discordant familes (Figure 31C, D), indicating that the escape of *PLS3* from XCI solely explains the expression difference in MNs.

CHD4 most likely acts as part of the NuRD complex (Marfella and Imbalzano, 2007; Tong et al., 1998; Wade et al., 1998; Xue et al., 1998; Zhang et al., 1998). The core proteins of NuRD contain several subunits, including the Metastasis-Associated proteins MTA1/2/3 (*MTA1* [MIM: 603526]; *MTA2* [MIM: 603497], and *MTA3* [MIM: 609050]), the Methyl-CpG Binding Domain proteins MBD2/3 (*MBD2* [MIM: 603547], and *MBD3* [MIM: 603573]), the Histone Deacetylases HDAC1/2 (*HDAC1* [MIM: 601241], and *HDAC2* [MIM: 605164]), the Retinoblastoma Binding Proteins RBBP4/7 (*RBBP4* [MIM: 602923], and *RBBP7* [MIM: 300825]), the GATA Zinc Finger Domain Containing proteins GATA2a/2b (*GATA2* [MIM: 137295]) as well as the motor subunits CHD3/4/5 (*CHD3* [MIM: 602120], and *CHD5* [MIM: 610771]). The MBD2/3 proteins have the capacity to selectively recognise methylated DNA (Hendrich and Bird, 1998).

Indeed, in a ChIP-seq data set in human breast cancer cell lines, MBD3 overlapped with the promoter of *PLS3,* indicating that CHD4/NuRD and MBD3 act together to activate or repress genes (Shimbo et al., 2013). CHD3 and CHD4 are responsible for the ATPase activity of the NuRD complex and harbour conserved PHD fingers, chromodomains and a DNA-binding domain (Woodage et al., 1997). Interestingly, there is also a link between CHD4 and neuromuscular disorders. Conditional knockout of CHD4 in mouse granule neurons impairs activity dependent genes and dendrite pruning (Yang et al., 2016), while overexpression of *PLS3* leads to a significant delay in axonal pruning (Ackermann et al., 2013).

The CHD4/NuRD complex was originally described as transcriptional silencer (Cai et al., 2014; Denslow and Wade, 2007; Marfella and Imbalzano, 2007). However, multiple studies have shown that CHD4/NuRD either activates or represses transcription depending on the cellular context (Hoffmeister et al., 2017; Yang et al., 2016). Overexpression of CHD4 in Flp-In™ T-REx™ 293 cells lead to differential expression of 115 genes with equal amounts of up- and downregulated genes (Hoffmeister et al., 2017). In p22 mouse cerebellum, CHD4 overlapped with 9842 transcription start sites (TSS). Approximately 96% of these TSS' were active promoters associated with the active histone mark H3K4me3 (Yang et al., 2016). Our dual-luciferase-promoter assays showed that CHD4 is indeed sufficient to activate transcription of a reporter gene (Figure 37A, B). A siRNA mediated knock-down of *CHD4* in EBV cells leads to a decrease in the expression of *PLS3*, while overexpression of CHD4 in HEK293T cells is associated with an increase of *PLS3* (Figure 32A, B). One mechanistic explanation would be that a change in the CHD4 level changes the stoichiometry between CHD4/NuRD and endogenous CHD3/NuRD. It is known that the subunit composition of NuRD is tissue specific, while the isoforms containing either CHD3 or CHD4 have a distinct nuclear localisation and a distinct set of target genes (Denslow and Wade, 2007; Hoffmeister et al., 2017).

6.1.5. CHD4 directly interacts with PLS3 promoter

It is not clear how CHD4/NuRD regulates transcription and it is sometimes stated that CHD4 misses DNA binding capacity (Arends et al., 2019). However, several ChIPseq and ChIP-qRT-PCR studies have shown that the protein has DNA binding capacity independent of other components of the NuRD complex or transcription factors (Bornelov et al., 2018). A ChIPseq in therapy-resistant tumour initiating glioblastoma

cells revealed a direct binding of CHD4 to the *PLS3* promoter (Chudnovsky et al., 2014). These results are in line with our ChIP-qRT-PCR experiments, in which we showed that the protein is able to directly bind to the promoter region of *PLS3* (Figure 35). Furthermore, a more recent study indicates that mutations in *CHD4* may disrupt DNA binding activity of the protein underlining the ability to directly interact with DNA *in vivo* (Novillo et al., 2021).

6.1.6. Transcription factors that interact with CHD4/NuRD

Promoter hypomethylation and overexpression of *PLS3* are hallmarks of several cancers (Bedford and van Helden, 1987; Brothman et al., 2005; de Capoa et al., 2003; Ehrlich, 2002; Gama-Sosa et al., 1983; Wolff et al., 2021). In SS, circulating $CD4^+$ T-cells show increased expression of *PLS3*, *TWIST1* and *GATA6* compared to normal $CD4^+$ T-cells (van Doorn et al., 2004; Wang et al., 2011; Wong et al., 2015). The promoter regions of all three genes have been found to be hypomethylated in SS $CD4^+$ T-cells, indicating an epigenetic mechanism of gene regulation (Wong et al., 2015). Interestingly, *TWIST1* and *GATA6* are known to be regulated by CHD4/NuRD (Mohd-Sarip et al., 2017; Tremblay et al., 2018). One limitation of our study is that we do not know which specific transcription factors interact with CHD4/NuRD. Most likely, CHD4/NuRD activates transcription of target genes by recruiting specific transcription factors (Wilczewski et al., 2018). A recent paper proposed that CHD4/NuRD interacts with the transcription factors GATA4, NKX2-5 and TBX5 to regulate embryonic heart development (Wilczewski et al., 2018).

The analysis of FB transcriptomes showed that *PLS3* is strongly expressed without difference between asymptomatic and SMA-affected individuals. This is in line with previous results (Oprea et al., 2008). It is unclear, which mechanisms control the expression of *PLS3* in FBs and if *PLS3* escapes XCI in FBs. The analysis of biallelic SNVs in our FB transcriptomes would not be informative, because the cell populations did not derive from single cell clones. The escape of the gene in asymptomatic individuals may be masked by the availability of CHD4/NuRD and specific transcription factors. Our data show that *CHD4* is strongly expressed in FBs of all siblings (Figure 31). If the transcription of *PLS3* mRNA reached its maximum capacity, differences in the expression would not been measurable.

Another puzzle is the expression of CHD4 itself. We have clearly shown that the expression of *CHD4* correlates with the expression of *PLS3* and *TCEAL4*, however, we do not understand the mechanisms that regulate the *CHD4* expression itself.

The transcriptional regulation of *PLS3* seems to be highly complex. Various epigenetic mechanisms, such as escape from XCI, the copy number of *DXZ4* and the epigenetic regulator CHD4/NuRD seem to influence the *PLS3* expression independently of each other or in combination. Over the last decade, the importance of PLS3 as genetic modifier was not only shown in neuromuscular disorders, it was also suggested as an important biomarker in several malignancies, as well as osteoporosis, osteoarthritis and in CAKUT (Fedou et al., 2021; Mählich et al., 2021; Makitie et al., 2020c; Xin et al., 2020). Unravelling the complex regulation of *PLS3* may improve our understanding of disease pathologies in this wide range of disorders. Especially in the cancer field, where PLS3 is involved in EMT, the gene regulation may be crucial to develop novel treatment strategies.

6.2. Part II: The impact of PLS3 on biomarkers of SMA

In our study, we analysed the plasma concentration of SMN and six putative biomarkers (COMP, DPP4, SPP1, CLEC3B, VTN, and AHSG) in an intermediate SMA mouse model, treated with presymptomatic low-dose *SMN*-ASO injections. Furthermore, we included mice with a human *PLS3* transgene. Only COMP and DPP4 showed high and SPP1 moderate correlation with the SMA phenotype. *PLS3* overexpression from a human transgene neither influenced the SMN level nor the six biomarkers, supporting the hypothesis that *PLS3* acts as an independent protective modifier of SMA.

6.2.1. Characteristics of biomarkers

Biomarkers must be easily accessible, or they will not be used in clinical practice. Spinal MNs are not accessible in patients. Spinal-cord fluid, muscle or skin tissue would be available; however, multiple testing would not be justifiable. Blood, saliva or urine of patients circumvent these disadvantages and is safe and cost-effective. However, the concentration of SMN in the hematopoietic system may not reflect the abundance in other tissues; in addition to that, spinal MNs are more affected by the abundance of SMN than other cell types (Coovert et al., 1997; Lefebvre et al., 1997).

SMA is a slowly progressive disease with a broad range of severities depending on the copy number of *SMN2*. Additionally, genetic modifiers, such as PLS3, influence the progression of the disease. More importantly, in the last decade multiple treatment options became available. With all these factors in mind, biomarkers for SMA must be able to reliably monitor the disease progression and treatment success at given time points in unrelated cohorts of patients. In addition to that, biomarkers must be specific for a given medical condition. Several biomarkers for SMA were linked to multiple other disease groups (Arellano-Garcia et al., 2010; Chen et al., 2017). Therefore, each biomarker should be critically evaluated and comorbidity should be taken into account.

6.2.2. Effect of sex and PLS3 overexpression

As previously discussed, increased levels of PLS3 lead to improved motor neuron axon outgrowth and NMJ functionality, extension of survival and motoric ability especially in low-dose *SMN*-ASO treated severe SMA mouse models (Ackermann et al., 2013; Alrafiah et al., 2018; Hosseinibarkooie et al., 2016; Kaifer et al., 2017). In our study,

murine *Pls3* was not expressed in blood of WT animals (Figure 39A, B). This goes well in line with the fact that *PLS3* is normally not expressed in the haematopoietic system but in all solid tissues (Lin et al., 1993). Furthermore, there were no differences in the expression of murine *Pls3* between male and female mice (Figure 39C). However, the overexpression of *PLS3* was achieved genetically from a human transgene inserted in the *Rosa26* locus on chromosome 6, while the murine *Pls3* is X-linked and located next to the murine *Dxz4* (Ackermann et al., 2013). The human *PLS3* transgene was ubiquitously expressed under the chicken beta promoter (Ackermann et al., 2013). This setup would resemble the situation in asymptomatic *SMN1*-deleted individuals (Riessland et al., 2017). In our data set, overexpression of *PLS3* did neither influence the blood concentration of SMN, nor the other six blood biomarkers. Based on our data, *PLS3* might work in other independent pathways, which might not be related to the regulation and release of these six putative biomarkers.

6.2.3. Effect of ASO-treatment

In the Taiwanese SMA mouse models, low-dose *SMN*-ASO treatment doubles the mean survival (Hosseinibarkooie et al., 2016). Interestingly, the low-dose subcutaneous *SMN*-ASO treatment did not affect the concentration of SMN and only mildly modulates the plasma levels of the six other plasma biomarkers. One explanation might be that the ASOs do not distribute equally *in vivo*. It has been found that ASOs distribute over the blood flow and accumulate in the liver, kidney, adipocytes, bone marrow and spleen, while haematopoietic cells show a low response to ASOs (Geary et al., 1997; Geary et al., 2003). Thus, the low-SMN-dose may counteract some of the failures in the development of the liver in severe SMA mouse models and explain the doubling in survival (Szunyogova et al., 2016). Additionally, SMN is an intracellular protein and it is highly likely that the plasma SMN protein originates from erythrocytes, lymphocytes or platelet cells that were lysed during the collection of plasma. Therefore, the amount of SMN in the plasma might not reflect the amount of SMN in the liver or in the spinal MNs. The intermediate SMN mouse model is an excellent model to study the rescuing effect of PLS3 overexpression. However, to generate our intermediate mouse model, we administered a low-dose of *SMN*-ASO (Hosseinibarkooie et al., 2016). In a clinical scenario, patients are treated with a higher dosage of *SMN*-ASOs, which could have a stronger impact in the protein levels of

biomarkers. Additional studies in human study cohorts are needed to evaluate the optimal time windows for biomarker analyses to optimise treatment control.

6.2.4. The SMN protein – a reliable blood biomarker

The inverse correlation between the *SMN2* copy number and the disease severity renders the SMN protein as an obvious biomarker candidate. In line with other biomarker studies, we did not find significant sex-specific differences in the blood concentrations of SMN or the other six putative plasma biomarkers (Crawford et al., 2012). The BforSMA study analysed the SMN plasma concentration, *SMN2* copy number and number of transcripts and found a correlation between SMN protein levels and the disease severity, similar to our results. However, SMN and the MHFMS were weakly correlated (Crawford et al., 2012). Of all tested biomarkers in our study, the plasma concentration of SMN was the only one that was stable over time. This renders SMN as probably the most important blood biomarker in our cohort. Furthermore, there were significant differences in the amount of SMN protein between the severely affected SMA mouse groups and the WT, however, the protein levels between HET and WT mice were largely overlapping. This result is in line with multiple studies, which have shown that SMN mRNA and protein concentrations in blood cannot be used to measure the severity of SMA (Crawford et al., 2012; Czech et al., 2015; Sumner et al., 2006; Tiziano et al., 2010). The overlap between the WT and HET mice might be explained by the high degree of noise in the data. The plasma proteins that were measured, originate from various cells of the hematopoietic system. Some proteins may be secreted into the extracellular compartments, but others may be released into the blood plasma during the sampling and processing of the plasma.

There was a large number of outliers in the data set, which we excluded by application of Tukey's method (Tukey, 1977). A decrease of data points always reduces the statistical power. Pooling of groups that showed no significant differences can bypass this loss of power but comes with some disadvantages: it artificially reduces the observed *P*-values, and information about individual biological variation gets lost. However, pooling of the samples allowed us to clearly discriminate SMA, HET and WT groups at both P10 and P21.

6.2.5. Abundancy of the SMA blood plasma biomarkers over time and comparison to other studies

There are multiple studies, which have analysed putative SMA plasma protein biomarkers comparable to our study. In the first study, plasma protein concentrations in 108 children with genetically confirmed SMA have been analysed and the results were correlated with MHFMS scores (Finkel et al., 2012). A second study identified and validated additional protein plasma biomarkers for SMA using an immunoassay (Kobayashi, 2013). In the third study plasma concentrations in a cohort of very young infants (age < six months) have been analysed. They used two different motor function tests for young infants, the Test for Infant Motor Performance Items (TIMPSI) and Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP-INTEND) (Kolb et al., 2016). In the most recent study, the levels of ten putative biomarkers have been measured at three different time points between *SMN*-ASO-treated SMAΔ7 and HET mice (Arnold et al., 2016). There are strong disagreements about the abundancy of putative SMA biomarkers over time between the various studies. In the following paragraphs, we will compare the results of the mentioned studies. The findings are summarised in table 11.

Table 11: Comparison to other plasma biomarker studies (modified from Strathmann et al., 2018)

	Arnold et al., 2016			**Finkel et al., 2012**	**Kolb et al., 2016**			**Strathmann et al., 2018**		**Overall trend**
Model	SMAΔ7 mouse Correlation with SMN			Infants Correlation with motor score	Infants Correlation with motor score			Taiwanese SMA mouse Correlation with SMN		
Time point/ measured variable	P12	P30	P90	MHFMS	Age	TIMPSI	CHOP-INTEND	P10	P21	
COMP	n.s.	n.s.	n.s.	↗	↘	↗	↗	↗	↗	↗
DPP4	↘	n.s.	n.s.	↗	n.s.	↗	n.s.	↗	↗	↗
SPP1	↘	↘	↗	↘	↘	n.s.	n.s.	↗	↘	→
CLEC3B	↗	↗	↗	n.s.	n.s.	↗	n.s.	↗	↗	↗
VTN	↗	↘	↗	↗	-	-	-	↗	n.s.	↗
AHSG	n.s.	n.s.	n.s.	↗	n.s.	n.s.	n.s.	↘	↘	→
↗ correlation			↘ anti-correlation		→ contradicting data				n.s. not significant	

6.2.5.1. COMP

COMP, SPP1 and CLEC3B are associated with human disorders that affect bone and connective tissue. COMP belongs to the Thrombospondin protein family and is mainly located in cartilage, bone tissue and other connective tissues (Agarwal et al., 2012; Muller et al., 1998; Newton et al., 1994; Wang et al., 2010). Increased serum concentrations were found in patients with rapid hip joint destruction (Mansson et al., 1995). The serum level of COMP was lower in children with idiopathic scoliosis and high COMP was modestly correlated with high growth velocity (Gerdhem et al., 2015). Genetic variants are associated with pseudoachondroplasia (Briggs et al., 1995; Hecht et al., 1995). COMP was suggested as biomarker for disease progression of rheumatoid arthritis, osteoarthritis and liver fibrosis, and early cartilage lesions in the knee (El Defrawy et al., 2016; Jiao et al., 2016; Kluzek et al., 2015; Zachou et al., 2017). In a recent study, COMP was used as a biomarker for knee osteoarthritis. In that study, 42 women with osteoarthritis in different stages were treated with a well-rounded exercise program over a period of 12 weeks leading to a significant increase of plasma COMP (Azukizawa et al., 2018). Paediatric SMA patients are reported to have a low bone mineral density and femur fractures are highly prevalent in all SMA subtypes (Khatri et al., 2008; Wasserman et al., 2017). This might explain the correlation of COMP plasma levels with SMA disease severity in our data set, as severely affected SMA mice are immobile. Interestingly, similar to PLS3, increased expression of COMP is associated with a poor prognosis in colorectal cancer (Nfonsam et al., 2019).

COMP showed a positive correlation with SMN at both time points in our study, while it has been stated as non-responsive to SMN in the SMNΔ7 mouse model (Arnold et al., 2016). However, COMP has been positively correlated to all three motor scores (Finkel et al., 2012; Kolb et al., 2016).

6.2.5.2. DPP4

DPP4 and AHSG were both suggested as biomarkers for non-alcoholic fatty liver disease (Lebensztejn et al., 2014; Zheng et al., 2017).

DPP4 was positively correlated with SMN at both time points in our study (Figure 46) and was also correlated with MHFMS and TIMPSI (Finkel et al., 2012; Kolb et al., 2016). In the previous studies, the protein levels were not correlated with the age of

infants, while the plasma levels slightly increased over time in our study. (Kolb et al., 2016). The SMNΔ7 mice showed a contradicting behaviour, as the DPP4 levels were anti-correlated with SMN at P12, while at later time points there have been no significant correlations with SMN (Arnold et al., 2016).

6.2.5.3. SPP1

SPP1 is a non-collagenous matrix protein that helps osteoclasts to migrate and attach to the mineral matrix of bone surfaces (Haylock and Nilsson, 2006). SPP1 is produced by a variety of cell types and is associated not only with bone remodelling but also with immune regulation, inflammation and vascularisation. Thus, expression of *SPP1* is correlated with tumorigenesis, progression and metastasis of several malignancies (Hao et al., 2017). Amongst others, misregulation in the level of SPP1 has been reported as possible biomarker for Duchenne Muscular Dystrophy and cardiovascular events (Kuraoka et al., 2016; van der Leeuw et al., 2016). In pancreatic ductal adenocarcinoma, high levels of *PLS3* are associated with poor overall survival. In addition to that, SPP1 is associated with *PLS3* expression in these patients (Xiong et al., 2021). *SPP1* mRNA was downregulated in bone marrow-derived stromal/pre-osteoblast cells of a Taiwanese SMA mouse model compared to the controls (Shanmugarajan et al., 2009). In line with that, we have found that plasma levels of SPP1 were decreased in SMA mice and HET mice at P10.

For SPP1, there are contradictory data. It is negatively correlated with the age of infants and with the MHFMS (Finkel et al., 2012; Kolb et al., 2016). In the current study, there is a positive correlation with SMN at P10 and a negative correlation at P21. In the SMNΔ7 mice it is the opposite (Arnold et al., 2016).

6.2.5.4. CLEC3B

CLEC3B is induced during the mineralisation phase of osteogenesis (Wewer et al., 1994). It was suggested as a prognostic biomarker for different types of cancer (Arellano-Garcia et al., 2010; Chen et al., 2017).

For CLEC3B we found a moderate correlation with SMN at both time points. Additionally, it is correlated with TIMPSI (Kolb et al., 2016). In SMAΔ7 mice CLEC3B has been significantly correlated with SMN only at the earliest time point (Arnold et al., 2016). One explanation for the strong disagreements between the studies might be

that several different model systems and measure values were used. In addition to that, the time points and number of measurements were different between the studies.

6.2.5.5. VTN

VTN has been positively correlated with MHFMS (Finkel et al., 2012). A correlation of VTN with SMN was found only at P10. In SMNΔ7 mice the results have been also inconsistent with a negative correlation at P30 and positive correlations at P12 and P30 (Arnold et al., 2016).

6.2.5.6. AHSG

Interestingly, AHSG is an important non-collagenous protein in bones involves in endochondral ossification (Brylka et al., 2017). It regulates bone remodelling and calcium metabolism and AHSG-deficient mice showed a shortened femoral bone phenotype (Brylka et al., 2017). The involvement of AHSG in bone development may explain the increased protein concentration in severely affected SMA animals. AHSG is mainly secreted from the liver and inhibits insulin receptor tyrosine kinase (Srinivas et al., 1993). The gene was also suggested as a biomarker for Alzheimer's disease, sepsis, cardiovascular mortality and stroke (Chen et al., 2013; Karampela et al., 2017; Kitamura et al., 2017; Stenvinkel et al., 2005).

AHSG is the only biomarker in our study with increased concentration in severely affected mice. The concentration was anti-correlated to SMN at both time points. In the SMNΔ7 model, AHSG did not correlate with SMN at any time point (Arnold et al., 2016). AHSG has been positively correlated with MHFMS and no correlation with age and motor scores in young infants has been found (Finkel et al., 2012; Kolb et al., 2016).

6.2.5.7. Correlations

At P10, COMP, VTN and CLEC3B showed the highest correlation to each other. According to the Reactome data base, all three proteins are localised in the extracellular region (Fabregat et al., 2016). COMP and VTN interact with several collagen types and are involved in the extracellular matrix organisation. This may indicate that they are affected by the same mechanistic pathways.

6.2.6. Other biomarkers for SMA

Biomarkers are of large interest and the number of biomarker studies is increasing every year. According to Pubmed, between 1970 and 2022 more than thousand papers

have been published that investigate different kinds of biomarkers to monitor the disease progression of SMA (Figure 60).

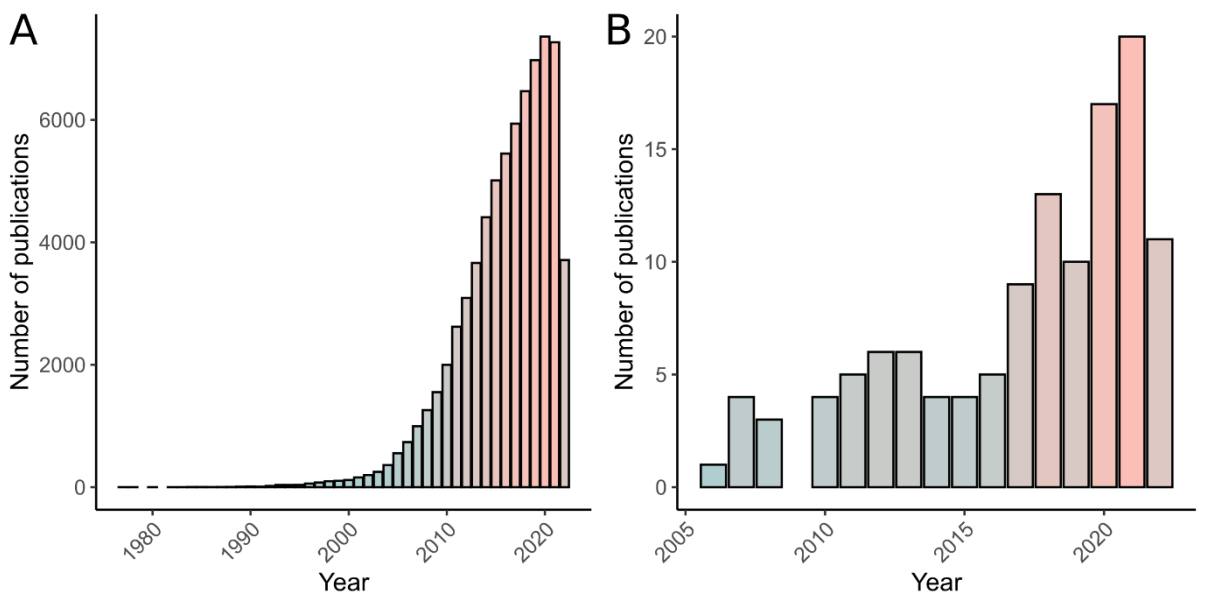


Figure 60: Rising number of biomarker studies. (A) A search query on Pubmed using the search terms "blood" AND "biomarkers" results in 61,219 studies, which have been published between 1970 and 2022. (B) A search for "biomarkers" AND "spinal muscular atrophy" results in approximately 106 hits.

Biomarkers for SMA can be categorised as biomarkers from solid tissues (muscle and skin biopsies, motor neurons) and blood biomarkers. Many biomarker studies concentrate on the analysis of plasma proteins (Arnold et al., 2016; Finkel et al., 2012; Kolb et al., 2016). However, other studies analysed miRNA or transcriptome data (Chen, 2020; Hawley et al., 2017; Magri et al., 2018; Viswambharan et al., 2017). A third method uses cell populations in peripheral blood to monitor SMA disease progression . In an early attempt, 23 putative protein biomarkers were measured in *Levator auris* longus muscle of a severe mouse model (Smn $^{-/-}$, SMN2$^{tg/tg}$). Calreticulin and GRP75/Mortalin were significantly increased in muscle biopsies of affected mice. Those results were reproduced in skin biopsies (Mutsaers et al., 2013). Later attempts focused on blood biomarkers as they are more accessible. Neurofilaments are scaffolding proteins that are found exclusively in axonal compartment of neuronal cells. They are released in the surrounding tissue and extracellular fluids after cell loss. Phosphorylation of heavy-chain neurofilaments (pNF-H) secures the filaments from degradation (Goldstein et al., 1987). Plasma levels of pNF-H are elevated up to 10-fold in SMA patients compared to age-matched controls (Darras et al., 2019). Treatment of infants with Nusinersen leads to a decrease of the pNF-H levels, which allows monitoring of the disease progress (Darras et al., 2019). However, in adult

SMA3 patients, the pNF-H levels did not change after treatment (Totzeck et al., 2019). Treatment with Nusinersen is associated with a decline of pNF-H (Darras et al., 2019). Despite contradicting results, neurofilaments seem to be a promising biomarker for SMA. Creatinine is a biomarker used to monitor muscle activity. SMA patients show lower levels of serum creatinine compared to controls (Alves et al., 2020). Another study used RNA sequencing of whole-blood samples from infant SMA patients and age-matched controls (Otsuki et al., 2018). Differential expression analysis revealed HSP70B as a novel biomarker. Further investigation has shown that the protein levels of HSP70B are associated with neurofilament levels in SMA infants (Eichelberger et al., 2021). One reason for the specific vulnerability of MNs to loss of SMN protein may be dysregulation of miRNA, as the protein is involved in miRNA biogenesis (Hawley et al., 2017).

In recent years, a rising number of research groups focused on circulating miRNAs as potential biomarkers for SMA (Chen, 2020; Hawley et al., 2017; Magri et al., 2018; Viswambharan et al., 2017). MiRNAs are small non-coding RNAs of about 22 nucleotides length and are involved in gene regulation. Early evidence for the usefulness of miRNAs as biomarkers for neurodevelopmental disorders came from a study in which the miRNA biogenesis pathway was disrupted by loss of *Dicer1* in mouse spinal motor neurons. This led to an SMA-like phenotype and a downregulation of miR-9, a miRNA that is known to regulate axonal extension and branching (Haramati et al., 2010). Next, altered levels of miR-9, miR-206 and miR-132 have been found in muscle and spinal cord tissue of SMA mice, as well as serum samples of SMA patients compared to healthy controls (Catapano et al., 2016). Next, miR-133a, miR-133b and miR-1 were identified as relevant miRNA biomarkers for SMA (Bonanno et al., 2020). The concentration of miR-133a, miR-133b, miR-206, miR-1 and miR-16 was measured in a cohort of 21 pediatric SMA patients (SMA2 and SMA3) before and after six months of Nusinersen treatment. A significant decrease was found for miR-133a, miR133b, and miR1 upon Nusinersen treatment, while miR133a reduction was found as predictor of treatment outcome (Bonanno et al., 2020). Another study measured miRNA levels in muscle biopsies of patients and controls and found over 100 differentially expressed miRNAs. Three of those miRNAs were significantly upregulated in serum of patients, namely miR-181a, miRNA-324, and miR-451a (Bonanno et al., 2020). Two other studies focussed on the measurement of miRNAs in cerebral spinal fluid. They found a general downregulation of miRNAs, however, miR-146a was upregulated in SMA

condition (Welby et al., 2022). MiR-206 and miR133a levels in cerebral spinal fluid can be used to predict the clinical response to Nusinersen treatment in mild type SMA (Magen et al., 2022). Recently, seven more miRNAs were identified, which are dysregulated in blood of SMA patients and respond to Nusinersen treatment, namely miR-107, miR-142-5p, miR335-5p, miR-423-3p, miR-660-5p, miR-378a-3p and miR-23a-3p (Zaharieva et al., 2022). A very different approach is the measurement of SMN levels in circulating cell populations as biomarker for SMA. CD3+, CD19+, and CD33++ cells in peripheral blood are significantly reduced in SMA condition (Otsuki et al., 2018).

6.4. Final conclusions/ future perspectives

Taken together, SMN, COMP and DPP4 showed the best performance as SMA biomarkers over time. All SMA biomarkers, except for SMN, are associated to multiple disease types and were suggested as biomarkers. Furthermore, SMN was the only plasma protein with stable concentration over time. Our results indicate that a reliable disease monitoring using plasma protein biomarkers for SMA is possible. However, to achieve optimal disease monitoring of SMA, we suggest to use a panel of multiple plasma proteins as biomarkers at the same time. The number of proteins that are necessary to ensure accurate measurements depends on the degree of correlation and co-regulation between them (Wang, 2011). The informative value of two or more biomarkers that share the same pathway is limited and addition of more proteins will not increase the predictive power substantially (Wang, 2011). One way to increase the informativeness would be to combine multiple biomarker approaches; for example, to combine plasma protein evaluation with a measurement of miRNA biomarkers as well as evaluation of phosphorylated neurofilament. However, this setup would increase both the costs and the effort. A biomarker assay panel, such as SMA-MAP seems to be a logic alternative (Kobayashi et al., 2013). This panel consists of 27 markers. It could be beneficial to include more biomarker candidates from the BforSMA pilot study, which identified 97 plasma proteins, 59 plasma metabolites and 44 urine metabolites that correlated with MHFMS scores (Crawford et al., 2012; Finkel et al., 2012). One idea to improve the diagnostic value of such a panel could be the implantation of computer-based analytical methods, such as deep-learning. Deep-learning methods are already in use in genetic counselling and are for example able to recognise genetic syndromes from facial dysmorphologies (Pantel et al., 2020). We propose that such

an algorithm could improve SMA diagnostics by evaluation of data from protein panels. However, a large amount of data is needed to train such an artificial intelligence.

Current SMA biomarkers will probably not replace the genetic diagnosis by established methods. However, they can be a useful tool for fast and reliable disease and treatment monitoring. There are disagreements about the correlation between the biomarkers and SMN between several studies, which might be explained by the use of different model organisms and the time points used in the various studies. Further studies are needed to evaluate these biomarkers in large cohorts of patients for final clinical application.

7. Materials and methods

7.1. Materials

7.1.1. Primers

Primer	Sequence	length [bp]	GC [%]	Melt [°C]	$T_{(anneal)}$ [°C]	cDNA [bp]
***PLS3* fwd**	GAACGTTGAGTGAAGCTGGA	20	50	57.3	63	358
***PLS3* rev**	TTGCCACTCTTCACAAGGTC	20	50	57.3		
***HPRT* fwd**	GCTATTGTAATGACCAGTCAACAGGGGAC	29	48.3	66.7	60	168
***HPRT* rev**	CCTTGACCATCTTTGGATTATACTGCC	27	44.4	63.4		
***TCEAL4* rev**	GGCTTCAAACCAAGGAAAGATGGAAAATG	29	41	58.7	60	632
***TCEAL4* fwd**	ATGCCTGCCAGGGACACTACACAT	34	54	59.1		
***DANT2* fwd**	ATCAGTAGACTGTGAGTTGTACATGGGCG AGA	32	47	63.1	63	210
***DANT2* rev**	CAGTCTGAGGAAGGTCAGTTGAAGTCCCT A	30	50	63		
***SLITRK4* fwd (FAM-tagged)**	GCACACAAGCAGTCCTTCCT	20	55	53.8	57	400
***SLITRK4* rev**	TGGCTTCTTGGTTGCTCTCT	20	50	51.8		
***CHD4* fwd**	CAAGGATGGTGGGGAACTCTG	21	57	56.3	60	362
***CHD4* rev**	GCTGATAGTTTCGGAACATCACCTGACA	28	46	59.9		
***MORF4L2* fwd**	GTTAATGTATAACAGGGCTTATGTTTCA	28	32	54.1	60	245
***MORF4L2* rev**	TTTTGTCACATAGCATTAACACATATTAG	29	28	53		

ChIP primers	Sequence	length [bp]	GC [%]	Melt [°C]	$T_{(anneal)}$ [°C]	cDNA [bp]
***PLS3* promoter fwd**	GTTGACAGAACTTTCCAAGCACACAT	26	42	56.4	56	300
***PLS3* promoter rev**	AATTTGGAGACATGTTACTACTCCC	25	40	54.4		
***TCEAL4* fwd**	GTGAGAAAGTGATATGGTTTGGGTCTGT	28	43	58.5	58	430
***TCEAL4* rev**	TCCTGTCCATTCTTGCACTGCTACAAA	27	44	58.2		

7.1.2. Laboratory equipment

Appliance	**Name**	**Manufacturer**
Analytical balance	Adventurer	Ohaus
Autoclave	VX-150	Systec
Bacterial incubator (shaking)	Innova 4230	New Brunswick
Binocular	DMIL	Leica
Bioruptor	Pico Sonification	Diagenode
Block heater	Thermo Shaker Incubator MSC-100	Universal Labortechnik
Cell counting chamber	Improved Neubauer chamber	Carl Roth
Combing device	FiberComb mol. combing system	Genomic Vision
DNA extraction for mol. combing	Fiber Prep DNA Extraction Kit	Genomic Vision
Fine balance	La314i	VWR
Fluorescence capillary electrophoresis	ABI3500 genetic analyser	Applied Biosystems
Freezer (-4°C/ -20°C)	Premium	Liebherr
Freezer (-80°C)	VIP plus	Panasonic
Gel electrophoresis chambers	Agagel SGE-020-02	Biometra CBS-Sci.
Geldocumentation	Chemidoc XRS Imaging System	BioRad
High-performance computer cluster	CHEOPS	
Imager	MSD 6000 Imager	Meso Scale Discovery
Incubator	Cell 150	Hera Cell
Liquid nitrogen store	MVEXLC 810HE	German Cryo
Luminometer	GloMax® 96	Promega
Magnetic rack	Dynamag2	Invitrogen
Magnetic stirrer	VMS-C70	VWR
Micro isolation chamber	Sentinel-cage-system	Bioscreen
Micro pipettes		Eppendorf
microliter syringe		Hamilton
Microwave	NN-E20JWM	Panasonic
pH-Meter	FiveEasy Plus	Mettler Toledo
Pipette-boy	Pipetboy2	Integra
Power supply for gel electrophoresis	PowerPac Basic	Bio Rad
Purification of DNA	MicroChIP DiaPure column	Diagenode
Qubit fluorometer	Qubit 4	Thermo Fisher Scientific
Real-time thermocycler	StepOnePlus RT-PCR system	Applied Biosciences
Scanner for Molecular Combing	FiberVision automated scanner	Genomic Vision
Sequencer (DNA)	HiSeqX	Illumina
Sequencer (RNA)	HiSeq2500	Illumina
Serum pipettes		Sarstedt
Sterile laminar flow hood	KS12	Kendro
Table centrifuge	5427R	Eppendorf
Thermocycler	C1000Touch	Bio Rad
Ultracentrifuge	Avarti J20 XPI	Beckman Coulter
Vacuum Manifold	Vac-Man®	Promega
Vacuum pump	Welch®	Promega
Vortexer	Vortexer	VWR
Water bath	1083 shaking water bath	GFL

7.1.3. Chemicals, reagents and media

Appliance	**Manufacturer**
2-β-mercaptoethanol	Life Technologies
Agarose	Sigma-Aldrich
Alfazyme	PAA
Amphotericin B	Thermo Fisher Scientific
Ascorbic acid	Sigma-Aldrich
Chir99021	Stemgent
Core lysis buffer	
Cyclosporine	Merck
De-ionised formamide	Merck
DMEM (high glucose)	Thermo Fisher Scientific
DMEM/F12	Life Technologies
DMSO	Sigma
DNA purple ladder plus	Thermo Fisher Scientific
Ethanol	Applichem
FBS	Sigma
GelRed Nucleic Acid Stain	Biotium
Glycine	Applichem
Isopropyl alcohol	Applichem
Laminin	Invitrogen
LDN 193189	Stemgent
Loading dye	NEB
mTeSR™1 medium	STEMCELL Technologies
N2 Supplement	Life Technologies
NaCl	Sigma
Nuclease-free H_2O	Lonza
Opti-MEM	Gibco
PBS	Gibco
Penicillin	Thermo Fisher Scientific
Poly-ornithin	Sigma-Aldrich
Proteinase K	Applichem
ROCK inhibitor Y-27632	Tocris Bioscience
RPMI 1640	Thermo Fisher Scientific
SB431542	Tocris
SDS (10%)	Applichem
SMN-ASO	IONIS Pharmaceuticals
Stem Pro Accutase	Thermo Fisher
Streptomycin	Thermo Fisher Scientific
TBE buffer	Applichem
TE^{-4} buffer	
Trypsin EDTa	Sigma
Y-27632	STEMCELL Technologies

Fresh blood lysis buffer (500 ml)

- NH_4Cl (1M) 77.5 ml
- $KHCO_3$ (10 mM) 5 ml
- EDTA (0.1 mM) 100 µl
- Fill up with H_2O and adjust to pH 7.4

Core lysis buffer (1000 ml)

- Tris (10 mM, pH 8.0) 10 ml
- NaCl (400 mM) 80 ml
- Na_2 EDTA 4 ml
- Fill up with H_2O and adjust to pH 8.2

TE^{-4} buffer (100 ml)

- Tris (10 mM, pH 8.0) 1 ml
- Na_2 EDTA (0.1mM, pH 8.0) 20 µl
- Autoclave

7.1.4. Kits

Appliance	**Manufacturer**	**Method**
DNeasy Blood & Tissue Kit	Qiagen	DNA extraction
Dual-Luciferase Reporter Assay System	Promega	Dual-luciferase-reporter assay
iDeal ChIP-qPCR kit	Diagenode	ChIP-qRT-PCR
Lipofectamine 3000	Thermo Fisher Scientific	Transfection
Monarch Total RNA Miniprep Kit	NEB	RNA extraction
Multiplex PCR kit	Qiagen	PCR
Purelink RNA Mini Kit	Thermo Fisher Scientific	RNA extraction
PureYield™ Plasmid Midiprep System	Qiagen	Extraction of plasmid DNA
QIAshredder kit	Qiagen	Biopolymer shredding
QuantiTect Reverse Transctiption Kit	Qiagen	Reverse transcription of RNA into cDNA
Qubit dsDNA BR Assay Kit	Invitrogen	Quantification of DNA
Qubit RNA BR Assay Kit	Invitrogen	Quantification of RNA
RNeasy Blood & Tissue Kit	Qiagen	RNA extraction
RNEasy Kit	Qiagen	RNA extraction
SMN ECL immunoassay	PharmOptima	Measurement of SMN
SybrGreen Master Mix	Applied Biosciences	qRT-PCR
TruSeq DNA library Prep Kit	Illumina	Library preparation for NGS
TruSeq RNA library Prep Kit	Illumina	Library preparation for NGS

7.1.5. Enzymes

Enzymes	**Catalog #**	**Manufacturer**
DNAseI	M0303S	NEB
*Hpa*II	ER0511	Thermo Fisher Scientific
*Msp*I	ER0541	Thermo Fisher Scientific

7.1.6. siRNAs

siRNA	**Catalog #**	**Manufacturer**
CHD4 siRNA	ID 121307	Thermo Fisher Scientific
AllStars negative Control siRNA	ID: 1027281	Qiagen

7.1.7. Antibodies

ChIP-grade Antibody	**host**	**Catalog #**	**Manufacturer**
anti-CHD4	rabbit	ab72418	Abcam
anti-CTCF	rabbit	C15410210	Diagenode
anti-IgG	rabbit	C15410206	Diagenode

7.1.10. Vectors

Name	**Manufacturer**
pLV[Exp]-CMV>hCHD4[NM_001273.5]	VectorBuilder
pRP[Exp]-PLS3(1.0kb)>Luc2	VectorBuilder
pGL4.10-empty>Luc2	Promega
pRL-TK	Promega
pRP[Exp]-CMV>Luc2	VectorBuilder

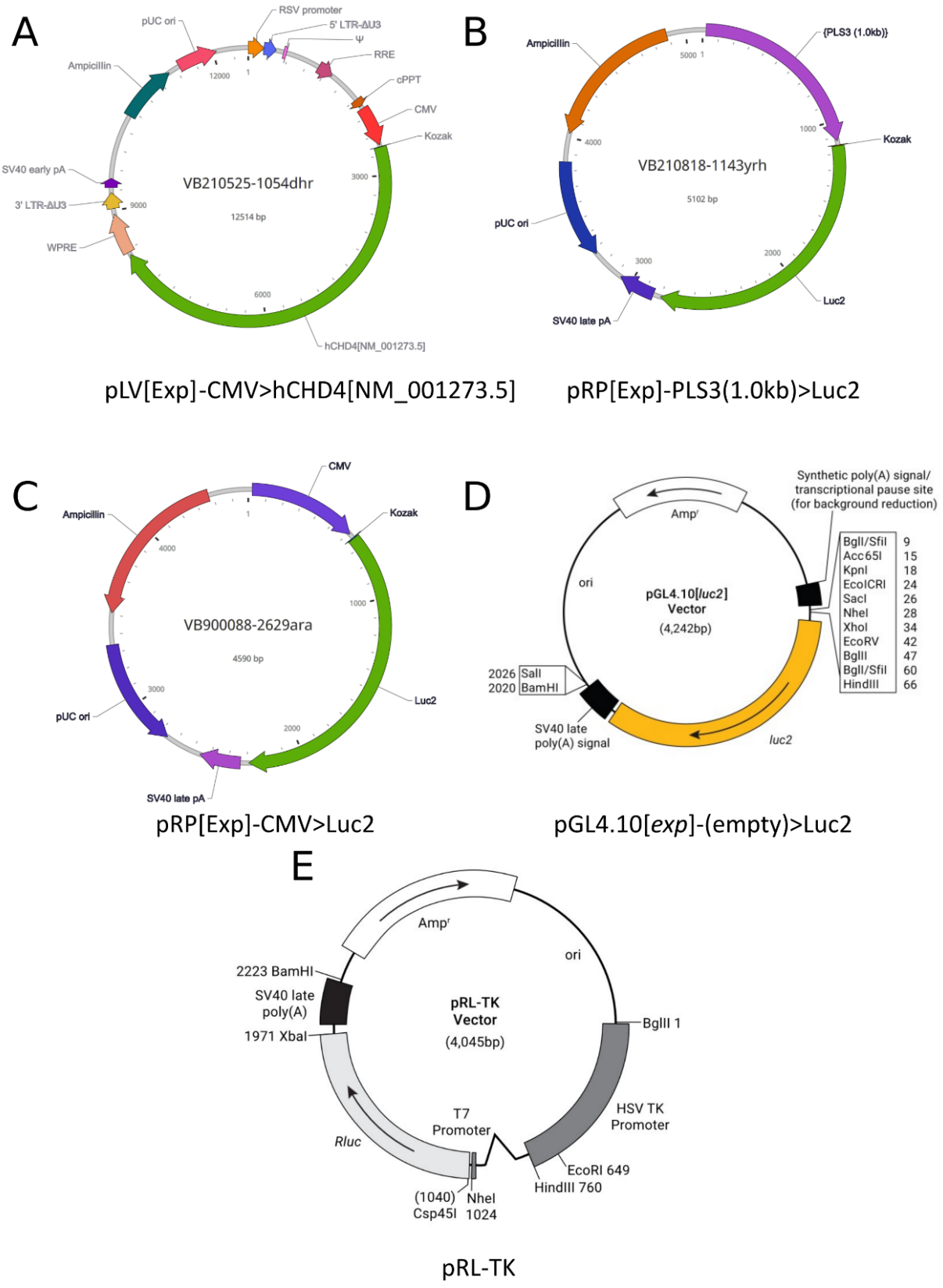


Figure 61: Vector maps of all plasmids that have been used for the dual-luciferase-promoter and the overexpression assays. (A) CHD4 overexpression vector under a *CMV* promoter.(B) Firefly-luciferase vector under the *PLS3* promoter (approximately 1 kb). (C) Firefly-luciferase vector under a *CMV* promoter as control. (D) Firefly-luciferase vector without promoter (empty control. (E) Renilla-luciferase vector as internal control for the dual-luciferase promoter assay.

7.1.11. Software

Appliance		
Microsoft Office 2016	Microsoft	Office Software
Endnote (Version 20)	Clarivate	Office Software
FiberStudio software	Genomic Vision	Molecular Combing
Fiji/ ImageJ	Fiji team	Image analysis
Inkscape	Inkscape Freedom Conservancy	Figure design
Windows 10	Microsoft	Operating system
Linux Mint Mate 19.2 64-bit	Linux Mark Institute	Operating system
Slurm Workload Manager		Operating system (CHEOPS)
R		
R-Studio		
R-packages		
ggplot2		
car		
Tximport		
DESeq2		
Dplyr		
biomaRt		
Keggrest		
Svglite		
EnhancedVolcano		
Tibble		
Ggpubr		
DescTools		
Bioinformatic tools		
NGSCheckMate		
Kallisto		
Samtools		
HISAT2		
FastQC		
DESeq2		
Cutadapt		
BCFTools		

7.1.13. Databases

Name	**Adress**
Ensembl	https://www.ensembl.org/
Gene cards	https://www.genecards.org/
Genecards	https://www.genecards.org/
gnomAD	https://gnomad.broadinstitute.org/
KEGG	https://www.genome.jp/kegg/
Ldlink	https://ldlink.nci.nih.gov/?tab=home
Mutation Taster	http://www.mutationtaster.org/
NCBI	https://www.ncbi.nlm.nih.gov/
Oligocalc	http://biotools.nubic.northwestern.edu/OligoCalc.html
OMIM	https://www.omim.org/
Pubmed	https://www.ncbi.nlm.nih.gov/pubmed/
SNPcheck V3	https://genetools.org/SNPCheck/snpcheck.htm
UCSC Genome Browser	https://genome.ucsc.edu/
UniProtKB	https://www.uniprot.org/help/uniprotkb
Reactome data base	http://www.reactome.org

7.2. Methods

7.2.1. Laboratory methods

7.2.1.1.Material derived from patients

Patient derived skin FBs have been established from three SMA3-affected and three asymptomatic siblings belonging to two unrelated SMA-discordant families (Heesen et al., 2016; Oprea et al., 2008). iPSCs from family 1 and 2 were previously generated and reported (Heesen et al., 2016). Epstein-Barr virus (EBV) transformed lymphoblastoid cells were generated from blood samples, including six asymptomatic females, 32 SMA-affected individuals with high and low expression levels of *PLS3* of both sexes, and six healthy controls as previously reported (Oprea et al., 2008) (Table 1). Informed written consent was obtained from patients, caregivers and family members according to the Declaration of Helsinki, and the study was approved by the ethics committee of the University Hospital of Cologne under the approval numbers 04-138 and 13-022.

7.2.1.2. Maintenance of EBV, HEK293T and fibroblast cell lines

EBV cells were maintained with RPMI 1640 medium with 20% FBS (Sigma), 1% Penicillin and Streptomycin (Thermo Fisher Scientific), and 6.25 µg/ml Amphotericin B (Thermo Fisher Scientific). The cells were cultured at 37°C and 5% CO2. Medium was added every 3 to 4 days. HEK293T cells and FBs were maintained in Dulbecco's Modified Eagle Medium (DMEM, high glucose) (Gibco) with 10% FBS (Sigma), 1% Penicillin/Streptomycin (Thermo Fisher Scientific), and 6.25 µg/ml Amphotericin B (Thermo Fisher Scientific). The medium was changed every two to three days.
When fibroblasts are 70-75% confluent cells are washed with PBS (Gibco) and split by incubation with 5 ml trypsin/EDTA (Sigma) for three to eight minutes at 37 °C. After detaching of the cells, the tyrosination was stopped by adding 5 ml fresh medium. Cells are collected and centrifuged for 3 minutes at 1,200 rpm. The pellets were resuspended in fresh DMEM (high glucose) medium (Thermo Fisher Scientific).

7.2.1.3. Generation of EBV cells

EBV cells were generated by transformation of lymphoblastoid cells from human blood. The blood samples were mixed with EDTA (1.5 ml/ ml blood) as an anticoagulant. A volume of 2.5 mL from children or 5 mL blood from adult individuals was used

respectively. The blood was transferred into a 50 mL Falcon tube and filled-up with sterile fresh blood lysis buffer to 50 ml. The samples were incubated for 10 minutes at room temperature and then centrifuged at 1,500 rpm for 10 minutes. The whole procedure was repeated a second time. After that, red blood cells at the edge of the white cell pellet were removed with a pipette tip. Next, 2 ml of an EBV suspension was added to the cell pellet, which was then incubated for 1h at 37°C. After that, 10 µl cyclosporine (1 mg/ ml) (Merck) and 3 ml RPMI 1640 medium (Gibco) were transferred into a T25 cell culture flask and the cell suspension was added. The cells were incubated at 37°C and the transformation process was examined daily.

7.2.1.4. Motor neuron differentiation from iPSCs

The iPSCs were grown on Matrigel™-coated cell culture plates in mTeSR™1 medium (STEMCELL Technologies). The medium was changed on a daily basis. The cells were split by treatment with Alfazyme (PAA) at a confluence of about 70% to 75%. The medium was supplemented with 10 µM ROCK inhibitor Y-27632 (Tocris Bioscience) after plating to single cell survival. iPSCs were differentiated into spinal MNs in 384 well plates as previously described (Maury et al., 2015). Briefly, hiPSC were dissociated enzymatically using Stem Pro Accutase (Thermo Fisher) for 5 minutes and plated in 25 cm^2 flasks (Dutscher) or spinner flasks (Corning) at a density of $0.2x10^6$ cell/ml in a total volume of respectively 10 ml and 70 ml. Cells were incubated in suspension in a neural induction medium containing of DMEM/F12 and Neurobasal medium supplemented with B27 (Life Technologies), N2 Supplement (Life Technologies), Penicillin/ Streptomycin 0.1%, β -ME 0.1% (Life Technologies), ascorbic acid (0.5 µM, Sigma-Aldrich), Chir99021 (3 µM, Stemgent), SB431542 (20 µM, Tocris), LDN 193189 (0.2 µM, Stemgent) and Y-27632 (10 µM, Stemcell). After 2 days, caudalisation of the neural progenitors was obtained by addition of retinoic acid (RA 0.1 µM) and smoothened agonist (SAG, 0.5 µM). After 1 week in suspension, brain derived neurotrophic factor (BDNF, 10 ng/mL), glial-derived neurotrophic factor (GDNF, 10 ng/mL), and N-[N-(3,5-Difluorophenacetyl)-L-alanyl]-S- phenylglycine t-butyl ester (DAPT, 10 nM) were added in the medium without SB431542 and LDN193189. Between day 10 and 14, MN progenitors were converted into MNs. At this step (day 10) embryoid bodies were dissociated into single cells with Trypsin-EDTA and plated either in 384-well plates (Corning) or cell culture dish coated with poly-ornithin (20 µg/ml, Sigma-Aldrich) and 5 g/ml laminin (Invitrogen) into the same medium with Y-27632. MNs were obtained after 14 days of differentiation and the

differentiation efficiency was assessed by immunolabeling for Islet1 (ISL1, from Neuromics Ref GT15051) and Hb9 (from DHSB Ref 81.5C10).

7.2.1.5. Freezing and thawing of cells

For long-term cryopreservation, cells were detached according to the respective splitting protocol. The cells were centrifuged and resuspended in freezing medium. The freezing medium for fibroblasts consisted of 1ml 90% FBS (Sigma), 10% DMSO (Sigma, sterile filtered). The freezing medium for EBV cells consisted of 1ml 90% RPMI1640 (Gibco), 10% DMSO (Sigma, sterile filtered). Cells were frozen in 1 ml cryotubes at -80°C in a freezing container. For long-term storage the samples were kept in liquid N2. Cells were thawed by fast warming in a 37 °C water bath. Then, the cells were transferred in 15 ml Falcon tubes filled with 5 ml pre-warmed (37 °C) medium. The tubes were centrifuged at 1,200 rpm at 4°C for 5 minutes and the pellet was suspended in fresh pre-warmed medium.

7.2.1.6. Extraction of RNA

RNA for sequencing and qRT-PCR experiments was extracted from densely populated EBV, FB or spinal MN cultures using the RNeasy Blood & Tissue Kit (Qiagen) or the Monarch Total RNA Miniprep Kit (NEB) according to the instruction manual. During this process, DNA was digested using DNAseI and the RNA was solved in 50µl nuclease-free H_2O.

7.2.1.7. Transcriptome analysis

Approximately 1,200 ng RNA (20ng/µl in 60µl) was sent for RNA sequencing to deCODE genetics, Iceland. For each sibling of family 1, RNAseq of iPSC-derived spinal motor neurons was performed in three replicates. For family 2, two independent iPSC clones with three replicates each were generated. In addition to that, RNA sequencing was performed on RNA from 40 EBV cell lines from six asymptomatic females, 32 SMA-affected individuals with various expression of *PLS3* of both sexes, and six healthy controls, including all siblings from families 1 and 2. Finally, RNA from six fibroblast cell lines was sent to RNA sequencing including all siblings from families 1 and 2.

7.2.1.8. Sanger sequencing

Genetic variants were validated by Sanger sequencing. This method of DNA sequencing uses dideoxynucleotide triphosphates to terminate the elongation of a DNA strand. The dideoxynucleotide triphosphates are labelled with fluorescent dye that allow the identification of the nucleotide by capillary electrophoresis. The sequencing was performed by GATC Biotech. DNA was extracted from EBV or FB cells mixed with DNA primers following the instructions of GATC Biotech.

7.2.1.9. DNA extraction for whole-genome sequencing

Genomic DNA (gDNA) was extracted from FBs using the DNeasy Blood & Tissue Kit (Qiagen) in accordance with manufacturer's instructions. Approximately 1,200 ng DNA (20ng/µl in 60µl) was sent for DNA sequencing to deCODE genetics, Iceland. The library preparation was done using the TruSeq DNA library Prep Kit (Illumina). Paired-end sequencing by synthesis using 2x150 cycles of incorporation and imaging was performed on Illmina HiSeqX sequencers.

7.2.1.10. Extraction of DNA (Salting-out method)

Attempts to establish the methylation-sensitive X-inactivation assay failed using column-based extraction kits. Therefore, we isolated DNA from EBV cells using a modified salting-out procedure (Miller et al., 1988).

EBV cells were fed with RPMI 1640 medium (Gibco). On the next day, approximately 12 ml of densely populated EBV culture were transferred to a 15 ml Falcon tube and centrifuged for 10 minutes at 1,500 rpm. The supernatant was removed. Then, 1,250 µl fresh blood lysis buffer was added and the pellet was resuspended in Core lysis buffer. Afterwards, 82.5 µl 10% SDS (Applichem) and 250 µl Proteinase K (20 mg/ml in H_2O, Applichem) were added. The tubes were incubated over night at 37°C in a heating block. The next morning, 412 µl 6M NaCl solution (Applichem) was added and the tube was shaken to precipitate proteins. The tubes were centrifuged for 10 minutes at 4,000 rpm at room temperature and the supernatant was transferred into 15 ml Falcon tubes. 1,750 µl isopropyl alcohol were added and the tube was shaken lightly. The tubes were centrifuged for 10 minutes at 4,000 rpm at room temperature, the supernatant was removed and the tubes were placed inverted on a tissue to remove the liquid. Finally, 100 µl of TE^{-4} buffer were added. The DNA pellet was resolved by

pipetting and transferred into a pre-labelled microcentrifuge tube. The DNA was stored at 4°C in the fridge or at -20°C in the freezer for long-term storage.

7.2.1.11. Qubit and Nanodrop

The concentrations of both DNA and RNA were quantified using Qubit fluorometer (Thermo Fisher). The Qubit dsDNA BR Assay Kit and the Qubit RNA BR Assay Kit were used in accordance to the manufacturer's user guides, respectively. Since both protocols are based on the same principle, only instructions for the DNA kit are given here. A master mix with 1 volume for each sample plus 2 volumes for the two standards was prepared. One volume of master mix contained 1 µl DNA BR reagent and 199 µl of RNA BR buffer. The two standards were prepared by adding 10 µl of the standard solution to 190 µl of the master mix. For each sample, 1 µl of DNA was added to 199 µl of master mix in 0.5 ml clear microcentrifuge tubes. The standards and samples were vortexed and incubated at room temperature for 2 minutes. After that, the standards were measured by Qubit to generate a calibration line. Then the samples were measured using the program DNA Broad Range Assay of the Qubit in accordance with the user manual.

7.2.1.12. qRT-PCR

In order to perform qRT-PCR, RNA has to be reversely transcribed into complementary DNA (cDNA). For that, the QuantiTect Reverse Transcription Kit (QIAGEN) was used in accordance with the manufacturer's instruction manual. Dilutions of 800 ng RNA in 12 µl RNAse-free H_2O were prepared. To eliminate contaminations of genomic DNA (gDNA), 2 µl of gDNA Wipeout buffer were added and the solution was incubated for 2 minutes at 42°C. A reverse transcription master mix was prepared on ice by adding 14 µl template RNA to 1 µl Quantiscript Reverse Transcriptase, 4 µl Quantitect RT buffer and 1 µl RT primer mix. The master mix was added to each sample, which was immediately incubated for 15 minutes at 42°C. The enzymes were inactivated by incubation of 3 minutes at 95°C.

qRT-PCR was performed using the SybrGreen Master Mix (Applied Biosciences) in accordance with the manufacturer's instruction manual with 20 ng cDNA per reaction. All experiments were performed using a StepOnePlus Real-Time PCR system (Applied Biosciences). *HPRT* was used as housekeeping gene. An internal reference sample was included to all 96-well plates to compare the expression levels of all measured

samples. All experiments were conducted in triplicates. The $2^{-\Delta\Delta C_T}$ method was used to calculate relative changes in gene expression (Livak and Schmittgen, 2001).

To quantify the expression levels of *Pls3* in mice, approximately 50–100 µL of blood, brain and muscle was collected from four WT mice and three female and three male heterozygous *PLS3* transgenic mice at P21. Isolation of total RNA from mouse tissues was performed using the Purelink RNA Mini Kit (Thermo Fisher Scientific) and the QIAshredder (Qiagen) kit according to the manufacturer's instructions. 400 ng of RNA was reversely transcribed into cDNA (QuantiTect Reverse Transctiption Kit) according to the manufacturer's instructions. The quantitative real time PCR was performed using PowerSybrGreen Master Mix (Thermo Fisher Scientific) on a StepOnePlus PCR System (Applied Biosystems). Specific primers for the *PLS3* transgene and endogenous *Pls3* were used as previously described (Ackermann et al., 2013). *Hprt* was used as a housekeeper gene. The $2^{-\Delta\Delta C_T}$ method was used to calculate relative changes in gene expression (Livak and Schmittgen, 2001).

7.2.1.13. Agarose gel electrophoresis

Nucleic acids were separated by size in an agarose gel matrix in an electric field, due to their negatively charged phosphate backbone. For a 1% gel, 0.8 g Agarose was solved in 80 ml TBE buffer (Applichem) and heated up in a microwave. 6 µl of GelRed (Biotium) was added and the solvent was poured into a gel chamber with well combs. After polymerisation of the gel, TBE buffer (Applichem) was added to the chamber and the combs were removed. The first gel slot of each row was loaded with 6 µl of DNA purple ladder plus. The DNA or RNA was mixed 1:1 with loading dye and pipetted to the gel slots. The nucleic acid fragments were separated in an electric field of 85 V for approximately 30 minutes. The bands were documented with the ChemiDoc XRS Imaging system (BioRad).

7.2.1.14. Transformation of DNA into chemically competent E. coli

Transformation of *Escherichia coli* was performed using a heat shock protocol. Chemically competent *E. coli* bacteria were thawed in ice and mixed with 2 µl plasma solution. The bacteria were incubated in ice for approximately 30 minutes and then heat shocked for 30 seconds in a water bath at 41°C. Then, the vials were placed back on ice. Next, 250 µl of SOC medium (Thermo Fisher Scientific) was added to each vial, which were then incubated for 1 hour at 37°C. The transformed bacteria were

transferred to an LB agar plate with 50 µg/ml Ampicillin and incubated over night at 37°C. Plasmids were extracted from transformed *E. coli* cultures using the PureYield™ Plasmid Midiprep System (Qiagen) in accordance with the instruction manual. DNA was solved in 500 µl RNAse-free water.

7.2.1.15. CHD4 siRNA knock-down assays

siRNA-mediated knock-down of *CHD4* was conducted in three male and three female EBV cell lines. 750,000 EBV cells were seeded into 24-well plates and transfected with 1 µM *CHD4* siRNA (Thermo Fisher Scientific, siRNA ID 121307) or a mock siRNA (AllStars negative Control siRNA, Qiagen, Cat. No. / ID: 1027281). Lipofectamine 3000 (Thermo Fisher Scientific) was used as transfection reagent according to the instruction manual and the transfection was carried out in six replicates. On the next day, we harvested the cells, extracted RNA, performed reverse transcription, and measured the expression of *PLS3* and *CHD4* by qRT-PCR. *HPRT* was used as housekeeping gene.

7.2.1.16. CHD4 overexpression assay

CHD4 was overexpressed in HEK293T cells using a *CHD4* expression plasmid. In the first experiment, 125,000 HEK293T cells were transfected with 25 ng, 50 ng, 75 ng or 100 ng of a CHD4 expression vector under a *CMV* promoter (pLV[Exp]-CMV>hCHD4[NM_001273.5], VectorBuilder). In a second experiment, 450,000 HEK293T cells were transfected with 250 ng, 500 ng, 1000 ng or 1500 ng of the CHD4 expression vector, respectively (for vector maps see section 7.1.10.). Lipofectamine 3000 (Thermo Fisher Scientific) was used as transfection reagent according to the instruction manual and the transfection was carried out in three to four replicates. At the next day, the cells were harvested, RNA was extracted and reverse transcribed to perform qRT-PCR. The expression of *CHD4* and *PLS3* was measured. *HPRT* was used as housekeeping gene.

7.2.1.17. Dual-promoter-luciferase assay

A fragment of 1,101 bp of the *PLS3* promoter (GRCh37/hg19 ChrX: 114,794,485-114,795,585) including the transcription start site was cloned into a Firefly-luciferase vector (pRP[Exp]-PLS3(1.0kb)>Luc2, VectorBuilder) (for vector maps see section 7.1.10.). The coding sequence of *CHD4* [NM_001273.5] was cloned into an vector under a *CMV* promoter (pLV[Exp]-CMV>hCHD4 , VectorBuilder). Next, 125,000

HEK293T cells were seeded into 24-well plates in 500 µl medium. On the next day, the cells were co-transfected with 20 ng of the Firefly-luciferase vector under the *PLS3* promoter, 10 ng of the Renilla-luciferase vector (pRL-TK, Promega) as internal reference and several concentrations (0 ng, 6.25 ng, 12 ng, 25 ng) of the *CHD4* overexpression vector using 0.75 µl Lipofectamine 3000 (Thermo Fisher Scientific) in accordance with the instruction manual. In addition to that, cells were co-transfected with 50 ng of a Firefly-luciferase vector under a *CMV* promoter (pRP[Exp]-CMV>Luc2, Promega) and 10 ng of the Renilla-luciferase vector (pRL-TK, Promega) as positive control. A negative control was co-transfected with 10 ng of a Firefly-luciferase vector without promoter (pGL-4.10>Luc2, Promega) and 10 ng of the Renilla-luciferase vector (pRL-TK, Promega). After 24 hours, cells were harvested, lysed for 30 minutes with 1x lysis buffer and the luciferase activity was measured using the Dual-Luciferase Reporter Assay System (Promega) on a GloMax® 96 luminometer (Promega). Three independent experiments with each four technical replicates per treatment group were performed.

7.2.1.18. Methylation-sensitive X-inactivation assays

DNA from EBV cells was isolated from 12 ml densely populated medium using a modified salting-out procedure (Miller et al., 1988). To assess the X-inactivation (XI-) status of the X-chromosomes in female EBV cells, a methylation-sensitive X-inactivation assay was performed (Bertelsen et al., 2011; Kiedrowski et al., 2011). 200 ng genomic DNA was digested at 37°C overnight either with 20 U methylation sensitive restriction enzyme *Hpa*II (Thermo Fisher Scientific) or with the methylation insensitive isoshizomer *Msp*I (Thermo Fisher Scientific) in a total reaction volume of 20 µl. Undigested DNA was used as control. Next, 5 µl of the DNA was amplified using FAM-tagged primers for the X-linked gene *SLITRK4* (see section 7.1.1. Primers) (Qiagen, Multiplex PCR kit) in accordance with the instruction manual (Mullis KB, 1986). The fragments were analysed by fluorescent capillary electrophoresis (Applied Biosystems, ABI3500 genetic analyser). The allele peak areas were calculated and the XI-status was calculated as previously described using equation 2 (Kiedrowski et al., 2011):

$$\frac{1}{A} = \frac{digested\ peak\ area 1}{digested\ peak\ area\ 2} * \frac{undigested\ peak\ area 1}{undigested\ peak\ area\ 2} + 1$$

$$A = proportion\ of\ allele\ 1\ on\ the\ active\ X - chromosome$$

Equation 2: Calculation of the XI-status

7.2.1.19. Cell counting

Cells were counted by application of 10 µl medium cell suspension into an Improved Neubauer (Carl Roth) counting chamber using a binocular (Leica). The average cell number of all four corner squares was calculated and the cell number determined as average number times 10^4 per ml.

7.2.1.20. Molecular Combing

This method uses multi-coloured DNA fiber probes to mark macrosatellite repeats on linear stretched genomic DNA molecules (Nguyen et al., 2011). Custom-made fiber probes targeting the *DXZ4* repeats as well as upstream and downstream regions were designed by Genomic Vision. Each individual repeat monomer was covered by two differently labelled (red and green) 1.1 kb fiber probes separated by a 0.4 kb gap. Additionally, a BAC (RP11-761E20, 179 kb) covered the region distally of *DXZ4* as telomeric probe visualised as a blue signal. A centromeric red probe was covered by the fosmid G2487005D12 and is located 35 kb from the *DXZ4* sequence. To perform the Molecular Combing, high molecular weight genomic DNA was extracted from 750,000 EBV cells using the Fiber Prep DNA Extraction Kit (Genomic Vision). The DNA was embedded in agarose plugs and treated with Proteinase K (Applichem) overnight at 50°C, melted at 68°C and digested with β-agarose (Sigma-Aldrich) at 42°C overnight. The DNA was gently mixed with 1.2 µl Combing buffer. Vinyl-silane coated coverslips were slowly plunged into the DNA solution at constant speed of 300 µm/s using the FiberComb Molecular Combing system (Genomic Vision). The coverslips were dried at 65°C for 2 hours. To hybridise the combed DNA with the fiber probes, the coverslips were de-hydrated by washing with increasing concentrations of 70%, 90% and 90% ethanol for 1 minute, respectively. After that, the coverslips were air-dried at room temperatures for 10 minutes protected from light. Fiber probes were mixed 1:1 with de-ionised formamide and incubated at 37°C for 30 minutes. Next, 20 µl of the fiber probe/formamide mix were transferred to a microscope slide and the combed coverslip was placed upside-down on the microscope slide avoiding air bubbles. The combed DNA and the fiber probes were co-denaturised for 5 minutes at 90°C in a humidified chamber and incubated 16 to 20 hours at 37°C. Finally, the coverslip was removed from the slide and washed 3 times with pre-warmed hybridisation buffer at 60°C. Next, the coverslip was placed upside-down on a microscope slide with 20 µl detection solution and incubated at 37°C in a humidified

chamber. The coverslip was removed and washed three times with detection washing buffer for 3 minutes, respectively. Then, the coverslip was washed with PBS (Gibco) for 3 minutes and de-hydrated by washing with increasing concentrations of 70%, 80% and 90% ethanol for 1 minute, respectively. The air-dried cover slips were scanned using the FiberVision automated scanner (Genomic Vision). The images were analysed using the FiberStudio software (Genomic Vision). The lengths of the *DXZ4* repeats and the upstream and downstream regions were marked manually and the copy number was calculated as the length of the DNA in kilobases divided by 3, as the *DXZ4* repeat monomer is 3 kb in size. Only signals that contained both upstream and downstream regions were counted as complete signals.

7.2.1.21. Chromatin immune precipitation

Chromatin immunoprecipitation (ChIP) was performed using the iDeal ChIP-qPCR kit (Diagenode) in accordance with the instruction manual. In short, for one immunoprecipitation $2.5x10^6$ EBV or HEK293T cells were fixed in PBS (Gibco) by adding 1/10 fixation buffer for 13 minutes at room temperature. Glycine was added in a proportion of 1/10 to stop the fixation for 5 minutes on a shaker. The cells were centrifuged at 500xg for 10 minutes at 4°C. Next, the cell pellet was washed twice with ice-cold PBS (Gibco). After centrifugation, 1 ml per million cells ice-cold lysis buffer (il1b) was added and the samples were incubated for 20 minutes at 4°C on a rotator. After centrifugation at 500xg at 4°C, the supernatant was discarded. The cells were re-suspended in 1 ml per million cells ice-cold lysis buffer il2, incubated for 10 minutes at 4°C and centrifuged at 500xg. Protease inhibitor cocktail was added to the shearing buffer iS1b in a 1:200 ratio and finally mixed with 0.1% SDS on ice. The cell pellets were mixed with 150 µl shearing buffer per 1.8 million cells. The DNA was sheared for 3-6 cycles using a Bioruptor Pico sonification device (Diagenode). To check the fragment size an aliquot of the sheared DNA was reverse crosslinked (incubation for 4 hours at 65°C), purified with a MicroChIP Diapure (Diagenode) column and analysed on a 1.5% agarose gel. An aliquot of 1% (2.5 µl) of the sheared DNA was kept aside as input. For one immunoprecipitation 30 µl of Protein A magnetic beads were washed 3 times with 1 ml ChIP buffer iC1b according to the manual. A ChIP reaction mix was prepared containing for one reaction 6 µl FBS (Sigma), 1.8 µl 200x protease inhibitor cocktail, 20 µl 5x iC1b buffer and 5 µg of the required ChIP-grade antibody (anti-CHD4 (rabbit), Abcam ab72418, anti-CTCF (rabbit), Diagenode, C15410210, anti-IgG (rabbit), Diagenode, C15410206). 70 µl of the ChIP reaction mix was added to each

sample and incubated at 4°C for 4-5 hours. The sheared DNA (250 µl) was added to the prepared beads and incubated on a rotator overnight at 4°C. The next day, the samples were centrifuged shortly and placed into a magnetic rack for 1 minute to allow the beads to be captured by the magnetic field, the supernatant was removed and the beads were washed with buffer iW1. The samples were incubated for 5 minutes on a rotator at 4°C. The washing steps were repeated as described for the wash buffers iW2 to iW4. DIB buffer was completed by adding 1 µl proteinase k to 100 µl. The IP samples were re-suspended in 100 µl of DIB buffer, while 97.5 µl were added to the input samples. The IP and input samples were incubated at 55°C for 15 minutes and at 100°C for another 15 minutes. Then the tubes were briefly spun down and placed in the magnetic rack for 1 minute. The supernatant, which contained the DNA, was purified using a MicroChIP DiaPure column (Diagenode). qRT-PCR was performed with input and immunoprecipitated DNA using primers directed to the promoters of *PLS3* and *TCEAL4* as well as *MB and H19* as controls (see methods section 7.1.1. Primers). The recovery was calculated as:

$$\%\, recovery = 2\text{^}(Ct\ input - 6.64 - Ct\ sample) \quad * 100\%$$

Equation 3: Calculation of the recovery

7.2.1.22. Mouse breeding

The Taiwanese SMA mouse model FVB.Cg-Tg (SMN2)2Hung Smn1^{tm1Hung}/J (stock number 005058) was obtained from Jackson Laboratory (Hsieh-Li et al., 2000). Both alleles of the murine *Smn* have been knocked out in this severely affected mouse model, while it contains two copies per allele of a human *SMN2* transgene (Smn$^{KO/KO}$; SMN2$^{tg/tg}$). Crossing *Smn*$^{KO/KO}$; SMN2$^{tg/tg}$ mice with *Smn*$^{KO/WT}$ in each F1 generation 50% SMA (*Smn*$^{KO/KO}$; *SMN2*$^{tg/0}$) and 50% HET (*Smn*$^{KO/WT}$; *SMN2*$^{tg/0}$) mice were obtained (Riessland et al., 2010). We further backcrossed these mice for more than seven generations to produce mice with congenic C57BL/6N background (Ackermann et al., 2013). *PLS3* overexpressing transgenic animals were used to generate SMA-*PLS3*het *(Smn*$^{KO/KO}$; *SMN2*$^{tg/0}$; *PLS3*$^{tg/0}$) and SMA-*PLS3*hom (*Smn*$^{KO/KO}$; *SMN2*$^{tg/0}$; *PLS3*$^{tg/tg}$) mice as well as HET-*PLS3*het (*Smn*$^{KO/WT}$; *SMN2*$^{tg/0}$; *PLS3*$^{tg/0}$) and HET-*PLS3*hom (*Smn*$^{KO/WT}$; *SMN2*$^{tg/0}$; *PLS3*$^{tg/tg}$) mice, all on C57BL/6N background. Neonatal mice were ear tagged, tail cut, and genotyped as previously described (Ackermann et al., 2013; Riessland et al., 2010). In addition, WT C57BL/6N mice were used as controls.

Mice were maintained at the SPF animal facility of the Institute for Genetics at the University of Cologne, kept under a 12 h light cycle, and given a 2918 and 2919 breeding diet (Harlan). Not more than four mice were kept in a micro isolation chamber (Sentinel-cage-system; each air ventilation system was independent of the laboratory air ventilation) with constant access to water and food. The litter consisted of FS 14 (Sniff) and was replaced once per week. All personnel in the animal facility were professional animal caretakers. A scoring system was used to evaluate the burden of each animal according to specific criteria as listed in table 12. Each animal was controlled and evaluated daily. If an animal reached more than 10 points (medium burden), it was controlled two times daily. Animals that reached a score of 20 or more (strong burden) were sacrificed immediately by cervical dislocation (P10) or CO_2 gassing (P21) followed by immediate cervical dislocation and extraction of whole blood into K_2EDTA-coated tubes to avoid coagulation. To increase the amount of blood, a cervical massage was performed. All animal procedures were conducted in accordance with European, national and institutional guidelines and protocols and were approved by the responsible government authority: Landesamt für Natur, Umwelt und Verbraucherschutz NRW. (Animal Licence: LANUV NRW under the reference numbers 84–02.05.20.12.120, 84–02.04.2014.A006 and 84–02.04.2015.A378).

Table12: Scoring system.

A. Common criteria	**Points**
I. Body weight	
• Not influenced	0
• Reduction of body weight < 5%	1
• Reduction of body weight 5-10%	5
• Reduction of body weight 11-19% or rather vertebra and pelvic bones palpable (BCS 2)	10
• Reduction of body weight ≥ 20% or rather vertebra and pelvic bones recognisable (BCS 1)	20
II. General state of health	
• Gleaming, smooth fur; clean body cavities, clear eyes	0
• Defects of the fur (decreased or increased body care)	1
• Blunt fur, untended body cavities, dull eyes, increased muscle tonus, minor scratches (no inflammation)	5
• Dirty fur, clotted/wet body cavities, abnormal posture, dull eyes, high muscle tonus, bent back, leaning head, medium scratches/ skin injuries (medium reddening, swelling, decelerated/defective wound healing)	10
• Cramps, paralysis, massive scratches/ skin injuries (medium reddening, swelling, defective wound healing)	20
III. Spontaneous behaviour	
• Normal behaviour (e.g.: sleeping, reaction to touching, curiosity, social contact/interaction)	0
• Minor departure from normal behaviour (e.g.: reduced reaction, normal food and water uptake)	1
• Extraordinary behaviour (e.g.: reduced or increased motoric abilities)	5
• Self isolation, lethargy, apathy, distinct hyperkinetic or rather. Stereotypic behaviour, disturbed coordination	10
• Pain yelling, Auto mutilation/-aggression	20
IV. Clinical indication	
• Normal respiration	0
• Minor departure from normal situation (e.g.: slightly increased respiration)	1
• Minor palpable temperature departure; clearly increased abdominal respiration	10
• Strong temperature departure (animal feels hot or cold); respiration sounds, dyspnoea, cyanoses	20
B. Specific experimental criteria	
Development of necrosis	
• No necrosis	0
• Necrosis of the tail	1
• Necrosis of the ears	5
• Necrosis of the hind-and forelimbs	20

Score	**Points**
No burden	0
Minor burden	1-9
Medium burden	10-19
Strong burden	≥ 20

7.2.1.23. ASO injection

On P2 and P3 six animals of each genotype were injected subcutaneously with 30 µg *SMN*-ASO, as previously described (Hosseinibarkooie et al., 2016). *SMN*-ASOs (IONIS Pharmaceuticals) were diluted in sterile PBS (Gibco) and the concentration was calculated using photometric density (AD260) (10 mg/mL working solution) (Hua et al., 2011). Subcutaneous injections with a microliter syringe (Hamilton) were administered as reported (Hua et al., 2011). This study included two cohorts of mice to allow endpoint whole blood and plasma at P10 in one cohort and longitudinal assessment out to 21 days in another cohort. Untreated SMA mice were only available for comparison at P10 due to a median survival of about 14 days.

7.2.1.24. SMN ECL immunoassay and plasma biomarker protein analysis

From each mouse 100–200 µl whole blood were collected at P10 or P21 directly into K_2EDTA-coated tubes to avoid coagulation. The tubes were mixed gently by inversion. Aliquots of 20–50 µL were collected and stored at -80°C and used for SMN protein analysis as previously described (Kobayashi et al., 2013).

PharmOptima carried out the SMN ECL immunoassays as previously described (Zaworski et al., 2016). Samples were vortexed and 5 µl were diluted in 795 µl dilution buffer to a final dilution of 1:160. Assay plates were read using an MSD 6000 Imager (Meso Scale Discovery). Data reduction from the SMN ECL Immunoassay was performed using software provided with the MSD 6000 Imager (Meso Scale Discovery). Whole blood SMN values were reported as ng SMN/ ml. Blood for plasma processing was transferred to 1.5 ml micro-centrifuge tubes and centrifuged at 800xg for 10 minutes at 4°C. After that, the supernatant was removed and transferred into a fresh 1.5 ml micro-centrifuge tube. An aliquot of 25 to 50 µl was stored at -80°C.

PharmOptima carried out the measurement of six SMA biomarkers COMP, DPP4, SPP1, CLEC3B, VTN, and AHSG using commercially available antibodies and calibration reagents and multiplex assay plates manufactured by Meso Scale Discovery as previously described (Arnold et al., 2016; Kobayashi et al., 2013). The 882 measurements of 126 blood samples were performed in duplicates. The mean value for each plasma biomarker analyte was used. Raw data of the plasma concentrations of SMN and the six plasma proteins at both time points are given in table 7.

7.2.1.25. Statistics

All statistical analyses were performed using R (version 4.1.2 (2021-11-01)) and RStudio (version2021.09.2 Build 382). All plots were generated using the ggplot2 package. To test if data were normally distributes, Shapiro-Wilk tests were performed (Shapiro and Wilk, 1965). Next, to test if the groups have equal variances, Levene tests were performed (Olkin, 1960).

In case of normally-distributed data, two groups were compared by Student's t-test's; otherwise, Wilcoxon rank sum tests were performed (Student, 1908; Wilcoxon, 1945). Multiple comparisons were conducted by ANOVA, followed by *post-hoc* Dunnett's tests or KW tests followed by *post-hoc* Dunn tests, if data were not normally distributed

(Dunn, 1961; Dunnett, 1955; Girden, 1992). Linear correlation was calculated as Spearman's rank correlation coefficient ρ. A *P*-value of less than 0.05 was considered as significant (*, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$) (Spearman, 1904).

We applied Tukey's method to exclude outliers of each genotype and biomarker from our raw biomarker data sets (Tukey, 1977) (Table 7). For each time point (P10 and P21), a separate cohort of animals was used. Therefore, we used an experimental design for unpaired data. In order to account for potentially non-normally distributed data, we applied nonparametric KW tests as *a priori* tests. Since the concentration of each of the seven proteins was measured in every single animal, the KW tests were corrected for multiple comparisons using the Bonferroni correction (Dunn, 1958). In cases of significant differences, we subsequently applied *post-hoc* Dunn tests to identify the groups that differ significantly from each other (Dunn, 1961). Dunn tests were corrected for multiple comparisons by use of the Holm correction (Holm, 1979). For correlation analysis of SMN concentrations with the six putative biomarkers, linear models were applied and the Spearman's Correlation Coefficient ρ were calculated.

7.2.2. Bioinformatics

7.2.2.1. Estimation of DXZ4 copy number by read depth analysis

The whole genome data sets were aligned to the human reference genome assembly GRCh37/ hg19 and the read depth for each position of the chromosomes 22 and X was counted by Samtools (Li et al., 2009). The average read depth for chromosome 22 and the *DXZ4* locus was calculated. The read depth of the variable region (chrX: 114,959,000-115,006,000) of *DXZ4* was normalised by the read depth of chromosome 22.

7.2.2.2. Transcriptome analysis

The library preparation was done using the TruSeq RNA library Prep Kit (Illumina). The sequencing was performed on a HiSeq2500 (Illumina) machine. For each sample, approximately 128,000,000 paired reads were generated. The transcriptome data was aligned to the human reference genome assembly GRCh37/hg19 using HISAT2 (Kim et al., 2019). In each sample, approximately 35,400 transcripts were annotated. Variants were called using BCFTOOLS (Danecek et al., 2021; Li, 2011). X-chromosomal biallelic SNVs were identified from vcf-files of motor neurons using the

following filter criteria: read depth ≥ 100, Phred-scaled quality score ≥30, and SNV ratio between 0.1 and 0.9.

7.2.2.3. Differential expression (Kallisto)

The differential expression analysis was performed using KALLISTO and DESeq2 (Bray et al., 2016; Love et al., 2014). Sample identity was validated using NGSCheckMate (Lee et al., 2017). A principal component analysis (built-in function of DESeq2) was performed to validate that samples with the same phenotype (*e.g.* asymptomatic) or biological replicates cluster together separately. The list of differentially expressed genes was filtered for p-values ≤ 0.001, a baseMean of ≥ 500 and an absolute log2FoldChange of ≥ 0.5.

$$RPKM = \frac{no.of\ reads\ mapped\ to\ a\ gene * 10^3 * 10^6}{Total\ number\ of\ mapped\ reads\ from\ given\ library * gene\ length\ in\ bp}$$

Equation 1: Calculation of RPKM for each given gene.

7.2.2.4. NGS Pipeline

Illumina TruSeq adapters were trimmed using Cutadapt (version 1.15) in paired-end mode (Martin, 2011). Trimmed reads shorter than 20 or low quality scores (<20) were filtered out.

```
Cutadapt –m 20 –q 20 –j 0 \
    -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC \
    -A AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCATT \
    -o <trimmed.1.fastq.gz> \
    -p <trimmed.2.fastq.gz> \
    <reads.1.fastq.gz> <reads.2.fastq.gz>
```

After that, the quality of the trimmed raw reads was checked using FastQC (version 0.11.5) (Andrews, 2010).

```
fastqc \
    –outdir <output directory>
    <trimmed.1.fastq.gz>
```

The trimmed fastq files were uploaded to the CHEOPS high performance computer cluster. The reference genomes of GRCh37 hg19 (release 92) and GRCh38 hg38 (release 99) were downloaded to CHEOPS, unzipped and the fastq-files were uploaded to CHEOPS.

```
# connect to the server:
> ssh -Y estrathm@cheops.rrz.uni-koeln.de

# upload all the reference genomes to the server:
# HG19:
> mkdir reference_hg19
> wget ftp://ftp.ensembl.org/pub/grch37/release-
92/fasta/homo_sapiens/dna/Homo_sapiens.GRCh37.dna.primary_assembly.fa.gz
> wget ftp://ftp.ensembl.org/pub/grch37/release-
92/gtf/homo_sapiens/Homo_sapiens.GRCh37.87.gtf.gz
> gunzip Homo_sapiens.GRCh37.cdna.all.fa.gz
> gunzip Homo_sapiens.GRCh37.dna.primary_assembly.fa.gz

# HG38 release 99 (16.1.2020)
> mkdir reference_hg38
> wget ftp://ftp.ensembl.org/pub/release-
99/fasta/homo_sapiens/dna/Homo_sapiens.GRCh38.dna.primary_assembly.fa.gz
> wget ftp://ftp.ensembl.org/pub/release-> 99/gtf/homo_sapiens/Homo_sapiens.GRCh38.99.gtf.gz
> gunzip Homo_sapiens.GRCh38.dna.primary_assembly.fa.gz
> gunzip Homo_sapiens.GRCh38.99.gtf.gz

# upload the fastq files to the server:
> scp hg-linux@134.95.36.125:"/<source_dir>/*fastq.gz" /<target_dir>/
```

After that, index files for HISAT2 were generated (Kim et al., 2019).

```
#!/bin/bash
#SBATCH --mem=12gb
#SBATCH --job-name=indexing
#SBATCH --time=05:00:00

hisat2-build /<path_to_reference>/Homo_sapiens.GRCh37.dna.primary_assembly.fa
/<path_to_reference>/Homo_sapiens.GRCh37.dna.primary_assembly
```

The alignment files were generated using following script.

```
#!/bin/bash
#SBATCH --cpus-per-task=12
#SBATCH --mem=12gb
#SBATCH --job-name=htseq
#SBATCH --time=10:00:00

sample=COL_sample
R1=HISEQ.R1.fastq.gz
R2=HISEQ.R2.fastq.gz

mkdir /<path_to_sampledir>/$sample/

echo 'start hisat2'
hisat2 -p 12 --dta -x Homo_sapiens.GRCh37.dna.primary_assembly -1 $R1 -2 $R2 | samtools view
-b -h > /<path_to_sampledir>/$sample/$sample.bam
echo 'hisat2 finished'

echo 'samtools sort started'
samtools sort /<path_to_sampledir>/$sample/$sample.bam >
/<path_to_sampledir>/$sample/$sample.sorted.bam
echo 'samtools sort finished'

echo 'indexing started'
# Samtools index (appr. 2 min)
samtools index -b /<path_to_sampledir>/$sample/$sample.sorted.bam
/<path_to_sampledir>/$sample/$sample.sorted.bai
echo 'indexing finished'

echo 'htseq started'
import pysam
python htseq-count -f bam -s no -r pos -i gene_id -m union
/<path_to_sampledir>/$sample/$sample.sorted.bam Homo_sapiens.GRCh37.87.gtf >
/<path_to_sampledir>/$sample/$sample.count.txt
echo 'htseq finished'
```

Quality check

The resulting bamfiles were downloaded from CHEOPS are a few quality checks were performed using the command line. First, samtools quickcheck was used to validate the integrity of the files. The program examines if the header and the end-of-file (EOF) is present (Li et al., 2009).

```
samtools quickcheck bamfile.bam
```

Next, variants were called using bcftools (Danecek et al., 2021; Li, 2011).

```
Bcftools mplieup –fasta-ref reference.fasta bamfile.bam | bcftools call bcftools view
bamfile.bcf | vcftutils.pl varFilter -> bamfile.vcf
```

Kallisto (DESeq2)

KALLISTO uses a pseudoalignment approach for rapid quantification of transcript abundance in RNAseq data (Bray et al., 2016). Firstly, the reference genomes GRCh37 hg19 (release 92) and GRCh38 hg38 (release 99 (16.1.2020)) were downloaded from the Ensembl ftp server, unzipped and a kallisto index file was generated. Kallisto uses the coding DNA reference file for better performance.

```
> wget ftp://ftp.ensembl.org/pub/grch37/release-92/fasta/homo_sapiens/cdna/Homo_sapiens.GRCh37.cdna.all.fa.gz
> gunzip Homo_sapiens.GRCh37.cdna.all.fa.gz
> wget ftp://ftp.ensembl.org/pub/release-99/fasta/homo_sapiens/cdna/Homo_sapiens.GRCh38.cdna.all.fa.gz
> gunzip Homo_sapiens.GRCh38.cdna.all.fa.gz
```

```
#!/bin/bash
# Kallisto indexing (k-mer length of 31):
kallisto index -i Homo_sapiens.GRCh37.cdna.all.idx Homo_sapiens.GRCh37.cdna.all.fa
```

After that, the kallisto quantification algorithm was used to generate the abundance files for each pair of fastq files.

```
#!/bin/bash

reference_idx=Homo_sapiens.GRCh37.cdna.all.idx
out=kallisto_hg19/COL0xyz
fastq1=forward.R1.fastq.gz
fastq2=reverse.R2.fastq.gz

mkdir $out

kallisto quant -i $reference_idx -o $out -b 50 --threads=4 $fastq1 $fastq2
```

The subsequent differential expression analysis, pathway annotation and co-expression analysis was performed using an R script as previously described (Bray et al., 2016; Love et al., 2014). In addition, expression plots for all significantly expressed genes, heat maps were generated and a principal component analysis was performed.

7.3. Tables and supplemental information

7.3.1. Tables

Table 1: List of patient materials. Given is the sample identifier (ID) and the phenotype and sex of each sample. Samples that belong to the same family are indexed by their family number (Fam#). The table indicates if DNA from fibroblasts was used to generate whole-genome data sets (WGS (FB)) and/or if RNA from EBV cells, fibroblasts or iPSC-induced motor neurons was used to generate RNA sequencing data.

Asymptomatic, females, homozygous *SMN1* deletion, 3-4 *SMN2* copies, high PLS3 expressers							
ID	**SMN2 CN**	**Fam#**	**PLS3 expression in EBV (qRT-PCR)**	**RNAseq (EBV)**	**RNAseq (MN)**	**RNAseq (FB)**	**WGS**
AS f1	3	Family 2	high	x	MN3.1/ MN3.1	FB3	x
AS f2	3	Family 2	high	x	MN4.1/ MN4.2	FB4	x
AS f3	3	Family 1	high	x	MN1	FB1	x
AS f4	3	Family 6	high	x			x
AS f5	4	Family 5	high	x			x
AS f6	3	Family 4	high	x			x
SMA, females, homozygous SMN1 deletion, 2-4 *SMN2* copies, low and high *PLS3* expressers							
ID	**SMN2 CN**	**Fam#**	**PLS3 expression in EBV (qRT-PCR)**	**RNAseq (EBV)**	**RNAseq (MN)**	**RNAseq (FB)**	**WGS**
SMA f1	3		high	x			x
SMA f2	3		high	x			x
SMA f3	3		high	x			x
SMA f4	2		low				
SMA f5	2		high	x			x
SMA f6	3		high	x			
SMA f7	4		low	x			
SMA f8	3		low	x			
SMA f9	4		low	x			
SMA f10	4		low	x			
SMA f11	2		low				
SMA f12	3			x			
SMA f13	2						
SMA f14	2			x			
SMA f15	2			x			x
SMA f16	3						
SMA f17	4	Family 3		x			x

SMA, male, *SMN1* homozygous deletions, 3-4 *SMN2* copies, low or high PLS3 expressers							
ID	**SMN2 CN**	**Fam#**	**PLS3 expression in EBV (qRT-PCR)**	**RNAseq (EBV)**	**RNAseq (MN)**	**RNAseq (FB)**	**WGS**
SMA m1	3		high				
SMA m2	3	Family 1	high	x	MN2	FB2	x
SMA m3	3	Family 6	low	x			x
SMA m4	4		low	x			
SMA m5	2		high				
SMA m6	3		high	x			
SMA m7	3		high				
SMA m8	3	Family 2	low	x	MN6.1/ MN6.2	FB6	x
SMA m9	4	Family 4		x			x
SMA m10	3	Family 2		x	MN5.1/ MN5.2	FB5	x
SMA m11	4			x			
SMA m12	4	Family 5		x			x
SMA m13	3			x			**x**
SMA m14	3			x			
SMA m15	3			x			
Healthy control							
ID	**SMN2 CN**	**Fam#**	**PLS3 expression in EBV**	**RNAseq (EBV)**	**RNAseq (MN)**	**RNAseq (FB)**	**WGS**
He f1				x			
He f2				x			
He f3				x			
He m1				x			
He m2	2			x			
He m3	2			x			
Additional samples for bioinformatic analyses of DXZ4 read depth							
ID	**SMN2 CN**	**Fam#**	**PLS3 expression in EBV**	**RNAseq (EBV)**	**RNAseq (MN)**	**RNAseq (FB)**	**WGS**
SMA f18			high				x
SMA f19	3		high				x
SMA f20	3		high				x
SMA f21			low				x
SMA f22			low				x
SMA f23			low				x
SMA f24			low				x
SMA m16			low				x
SMA m17			low				x
SMA m18			low				x
SMA m19			low				x
SMA m20			low				x
SMA m21			low				x

Table 2: Differentially expressed genes from motor neurons of 2 families. A total number of 80 genes was differentially expressed, including *PLS3*. Given are the gene name and the log2-fold change. A negative log2 fold-change indicates stronger expression in asymptomatics compared to SMA-affected siblings.

Gene	log_2 Fold-Change	Adjusted *P*-value	Chromosome
HS6ST2	-2.00383	1.20E-09	X
RP11-706O15.1	-1.56139	3.90E-20	X
TMEM164	-1.26662	9.51E-25	X
MID1	-1.25196	4.66E-06	X
AMOT	-1.19028	8.54E-10	X
FRMPD3	-1.18981	6.09E-12	X
NLGN4X	-1.16207	7.15E-08	X
NHS	-1.151	3.05E-12	X
FGF13	-1.14844	7.83E-09	X
GYG2	-1.1339	4.95E-16	X
PLS3	**-1.12314**	**7.34E-14**	**X**
NAP1L3	-1.12161	4.54E-15	X
PAK3	-1.05673	1.77E-07	X
ARSD	-1.03893	1.86E-17	X
KLHL13	-1.03073	2.65E-05	X
UBA1	-1.0111	7.82E-16	X
ZFX	-0.9761	2.13E-11	X
KLHL15	-0.9423	4.75E-10	X
LRCH2	-0.94042	4.12E-10	X
EIF2S3	-0.94011	5.59E-21	X
HDHD1	-0.93886	3.49E-19	X
AP1S2	-0.92934	5.80E-09	X
TMSB4X	-0.92813	6.04E-08	X
CTPS2	-0.9217	2.69E-29	X
TXLNG	-0.92082	1.95E-14	X
TCEAL3	-0.91467	4.91E-08	X
TSC22D3	-0.90346	3.26E-13	X
SMS	-0.89952	2.51E-07	X
PRPS1	-0.8904	1.21E-09	X
PRKX	-0.88344	5.92E-17	X
ATP11C	-0.86934	2.61E-06	X
EIF1AX	-0.86885	8.95E-07	X
RAP2C	-0.86405	5.66E-16	X
PSMD10	-0.86312	1.69E-15	X
CASK	-0.84482	1.27E-08	X

Gene	log_2 Fold-Change	Adjusted *P*-value	Chromosome
MORF4L2	**-0.84011**	**2.44E-15**	**X**
POLA1	-0.82537	8.50E-17	X
TAB3	-0.82448	0.000939	X
SYAP1	-0.8207	4.95E-19	X
RBBP7	-0.81384	4.03E-10	X
ALG13	-0.80065	1.42E-09	X
GPM6B	-0.79756	0.000795	X
CDK16	-0.76172	9.83E-10	X
KDM6A	-0.74476	7.53E-22	X
NGFRAP1	-0.73028	6.83E-07	X
BEX2	-0.72002	1.83E-07	X
TRAPPC2	-0.71338	9.60E-10	X
ZRSR2	-0.71199	0.000545	X
ZCCHC18	-0.71141	0.000231	X
PJA1	-0.70438	3.49E-05	X
USP11	-0.69716	4.51E-08	X
CXorf57	-0.69664	7.58E-05	X
TCEAL4	**-0.68071**	**0.000432**	**X**
WDR44	-0.66706	0.000422	X
FAM199X	-0.66472	1.44E-06	X
CD99L2	-0.64874	0.000611	X
ATP6AP2	-0.64863	0.000281	X
PNPLA4	-0.62686	0.000461	X
CHM	-0.60908	3.30E-06	X
PRRG1	-0.60313	0.000145	X
TBL1X	-0.55648	0.000891	X
NHLH2	-0.67111	0.00027	1
IFFO2	-0.66161	7.10E-05	1
DDAH1	0.726913	0.000378	1
PDZRN3	-2.26147	1.30E-06	3
ZNF595	1.83634	3.87E-10	4
PCDHA6	-1.93427	3.48E-06	5
UNC5A	-0.87381	0.000681	5
FSTL4	-0.81061	5.90E-05	5
HS3ST5	1.098801	0.000795	6
PLXNA4	-1.34327	0.000354	7
MTUS1	1.501163	1.67E-05	8
GRK5	0.686904	4.06E-09	10
CHST11	-0.58374	8.16E-06	12

Gene	log_2 Fold-Change	Adjusted *P*-value	Chromosome
HERC2P3	1.428443	0.000617	15
NETO2	-0.91176	0.00019	16
TMEM97	0.511835	3.50E-05	17
PEG3	9.296714	2.07E-14	19
BMP7	-0.56446	1.07E-05	20
NPTXR	-0.52322	0.000102	22

Table 3: List of escape genes identified by biallelic SNPs. Genes of interest are underlined.

Gene	Pos.	REF	ALT	Average Read Depth	SNP ratio	Bialleleic & differential expression
ARSD	2824377	T	C	102	0.78	✓
SYAP1	16779167	A	G	102	0.87	✓
TAB3	30873245	C	T	104	0.27	✓
CHM	85219021	T	C	112	0.40	✓
FRMPD3	106847032	G	A	113	0.77	✓
ALG13	110970565	C	T	117	0.63	✓
ATP11C	139015138	T	A	120	0.53	✓
CASK	41374397	G	A	131	0.51	✓
KLHL15	24005086	T	C	133	0.53	✓
KDM6A	44938563	G	A	135	0.52	✓
TBL1X	9686834	G	A	143	0.53	✓
GYG2	2777985	C	T	148	0.46	✓
GPM6B	13835096	T	C	150	0.46	✓
TMEM164	109419165	G	A	162	0.50	✓
NLGN4X	5809258	G	A	172	0.57	✓
HDHD1	6967175	G	A	174	0.45	✓
AMOT	112020692	A	T	182	0.58	✓
TCEAL4	**102842041**	**T**	**C**	**184**	**0.30**	✓
TRAPPC2	13730787	G	C	184	0.43	✓
RP11-706O15.1	3735732	G	A	188	0.55	✓
FAM199X	103437789	A	G	191	0.70	✓
CD99L2	149935100	T	C	201	0.72	✓
SMS	21958851	A	G	201	0.38	✓
PRKX	3524309	T	C	208	0.51	✓
ZRSR2	15838366	C	T	211	0.55	✓
HS6ST2	131760779	C	T	212	0.28	✓
BEX2	102564583	A	G	219	0.72	✓
EIF2S3	24075862	A	G	224	0.69	✓
FGF13	137714384	A	G	224	0.22	✓
PLS3	**114879399**	**T**	**C**	**225**	**0.39**	✓

RAP2C	131348189	A	G	230	0.77	✓
MORF4L2	**102930960**	**A**	**G**	**236**	**0.64**	✓
NAP1L3	92927634	G	C	236	0.59	✓
PJA1	68381912	G	A	241	0.48	✓
TMSB4X	12993286	C	T	298	0.73	✓
CDR1	139866477	T	C	100	0.86	
RPS6KA3	20171887	G	A	101	0.49	
SHROOM2	9915165	C	T	102	0.37	
MTMR1	149932521	C	G	102	0.52	
KLHL4	86887244	C	T	103	0.49	
CXorf40B	149100561	C	G	103	0.73	
RP1-23K20.2	130164980	G	C	103	0.86	
ZNF280C	129338766	G	A	105	0.73	
SMC1A	53449568	G	A	106	0.15	
RP11-706O15.3	3785654	G	A	108	0.44	
KAL1	8498979	G	A	109	0.53	
PHKA2	18972497	C	G	112	0.50	
DMRTC1B	72065722	C	T	112	0.54	
NUDT10	51075841	A	G	112	0.75	
CXorf56	118672671	T	C	114	0.18	
DACH2	86083173	C	A	114	0.65	
CXorf40A	148628709	G	C	115	0.63	
GPC4	132437337	G	A	116	0.66	
MOSPD1	134022913	A	C	117	0.49	
PIR	15415583	C	T	118	0.62	
PGK1	77384627	A	G	120	0.78	
EIF4BP7	110863752	G	C	123	0.34	
DDX26B	134713855	G	A	125	0.49	
CA5BP1	15693857	C	T	128	0.54	
LDOC1	140271268	G	C	132	0.81	
LAMP2	119571796	T	C	133	0.84	
BHLHB9	101975725	G	C	136	0.63	
MXRA5	3228891	A	G	137	0.40	
PCYT1B	24578227	A	G	137	0.53	
FUNDC1	44401328	G	A	137	0.76	
XIAP	123046338	A	G	139	0.35	
TMSB15B	103227906	A	G	139	0.61	
CHRDL1	109931856	T	C	141	0.55	
PCDH11X	91874407	T	C	150	0.62	

RLIM	73806379	G	A	152	0.61	
DOCK11	117819773	A	C	154	0.38	
FAM127C	134155904	G	T	158	0.15	
ARMCX3	100880857	G	A	158	0.18	
TMEM185A	148682039	T	C	160	0.75	
GPR64	19008280	G	A	161	0.43	
LINC00086	134557791	G	A	163	0.25	
CA5B	15782747	C	A	178	0.19	
PHF6	133547518	C	A	180	0.43	
JADE3	46920104	G	C	181	0.52	
CLCN4	10204537	G	A	188	0.86	
TMEM47	34648280	G	A	192	0.64	
USP9X	41091941	A	G	194	0.58	
ZBTB33	119389966	C	T	197	0.51	
MBTPS2	21861434	A	G	202	0.39	
IDS	148560581	G	A	204	0.27	
ACOT9	23722835	A	G	206	0.62	
PLXNA3	153696766	G	A	208	0.40	
BEX4	102471531	G	T	210	0.16	
TCEAL7	102585902	G	C	210	0.29	
EFNB1	68060503	C	T	215	0.56	
HMGB3	150156339	A	G	220	0.54	
AIFM1	129283520	A	G	222	0.53	
APOO	23858452	G	A	223	0.47	
SNX12	70280108	C	T	224	0.74	
FHL1	135293082	G	A	226	0.74	
TSPAN6	99884828	G	A	228	0.51	
RBMX	135961587	G	T	228	0.70	
DDX3X	41208310	A	C	230	0.76	
HTATSF1	135593629	C	T	231	0.26	
PGRMC1	118377608	C	T	232	0.28	
SEPT6	118763406	C	T	234	0.51	
GLRA2	14627144	T	C	236	0.69	
SAT1	23801311	C	T	237	0.72	
TSPAN7	38547602	A	T	238	0.50	
SMARCA1	128633717	T	C	239	0.32	
MAGED1	51644748	C	A	240	0.69	
MED14	40573157	G	A	244	0.61	

Table 4: Genes with SNVs identified in the 3 male siblings in motor neurons.

MN2	MN6	MN5	Cause
PPP2R3B	*PPP2R3B*	*PPP2R3B*	
SLC25A6	*SLC25A6*	*SLC25A6*	Several non-transcribed pseudogenes
ASMTL	*ASMTL*	*ASMTL*	
AKAP17A	*AKAP17A*	*AKAP17A*	
CD99	*CD99*	*CD99*	Pseudogene adjacent to locus
NUDT10	*NUDT10*		
PGK1	*PGK1*		3 pseudogenes
XIAP	*XIAP*		
IL3RA	*IL3RA*	*IL3RA*	Alternatively spliced isoforms
CXorf56	*CXorf56*		
EIF2S3	*EIF2S3*		Paralog on Y-chromosome
EIF4BP7	*EIF4BP7*		Pseudogene
LAMP2	*LAMP2*		
PHF6	*PHF6*		
CXorf40B	*CXorf40B*		

Table 5: Identification of possible transcription factors in EBV cells.

Gene	log2FoldChange	A
USP17L20	-6.28469	0.00011
MEST	-4.37305	4.62E-06
PLS3	**-3.59788**	**2.52E-05**
HPCAL4	-3.36375	0.000452
CADM3	-3.18514	0.000894
IQCG	-2.97058	0.000289
LRP3	-2.88601	1.20E-05
MYH14	-2.37815	1.82E-05
PRAME	-2.28872	5.66E-05
FHOD3	-2.23641	0.000271
ZFHX4	**-2.04352**	**0.00011**
IFIT1	2.606526	0.000132
PLAC8	2.660837	1.20E-05
TRPM2	3.245335	0.000117
IL8	3.85761	9.59E-08
C20orf197	4.149417	0.000542
RPS28	5.079282	2.22E-08
IGKV1-33	7.492359	0.000847
IGKV2-28	10.52474	0.00011

Table 6: List of samples than contain the genetic variant rs1639122.

Sample	**genotype**	***PLS3* status (EBV cells)**
AS f1	hom	high expresser
AS f2	hom	high expresser
AS f5	het	high expresser
AS f6	het	high expresser
He m2	het	no data
He m3	het	no data
SMA f3	het	high expresser
SMA f6	hom	high expresser
SMA f7	hom	low expresser
SMA f9	het	low expresser
SMA f10	hom	low expresser
SMA f12	het	no data
SMA f15	het	no data
SMA m2	het	high expresser
SMA m6	het	high expresser
SMA m8	hom	low expresser
SMA m10	het	no data
SMA m11	het	no data
SMA m12	het	no data
SMA m15	hom	no data

7.3.2. Data sets

Table 7: Measurements of plasma concentrations of SMN, COMP, DPP4, AHSG, SPP1, VTN and CLEC3B in 21 treatment groups. Given are the mouseID, the average plasma concentration, the sex, sample group, and the sampling day. It is indicated, if the measurement was excluded as outlier by Tukey's method.

MouseID	Biomarker	Mean	Sex	Group	Sampling day	Outlier
p10-173	AHSG	128.59	female	HET	p10	
p10-174	AHSG	139.36	female	HET	p10	
p10-187	AHSG	164.26	female	HET	p10	*
p10-128	AHSG	140.92	male	HET	p10	
p10-129	AHSG	133.38	male	HET	p10	
p10-130	AHSG	147.32	male	HET	p10	
p10-141	AHSG	174.49	female	HET:PLS3het	p10	
p10-142	AHSG	145.53	female	HET:PLS3het	p10	
p10-159	AHSG	165.58	female	HET:PLS3het	p10	
p10-149	AHSG	189.07	male	HET:PLS3het	p10	
p10-160	AHSG	152.35	male	HET:PLS3het	p10	
p10-172	AHSG	176.52	male	HET:PLS3het	p10	
p10-120	AHSG	96.67	female	HET:PLS3het+ASO	p10	
p10-157	AHSG	125.59	female	HET:PLS3het+ASO	p10	
p10-224	AHSG	205.72	female	HET:PLS3het+ASO	p10	*
p10-123	AHSG	133.49	male	HET:PLS3het+ASO	p10	
p10-158	AHSG	124.52	male	HET:PLS3het+ASO	p10	
p10-223	AHSG	85.97	male	HET:PLS3het+ASO	p10	
p10-148	AHSG	132.17	female	HET:PLS3hom	p10	
p10-181	AHSG	118.52	female	HET:PLS3hom	p10	
p10-194	AHSG	127.49	female	HET:PLS3hom	p10	
p10-150	AHSG	158.47	male	HET:PLS3hom	p10	
p10-152	AHSG	146.71	male	HET:PLS3hom	p10	
p10-178	AHSG	138.6	male	HET:PLS3hom	p10	
p10-122	AHSG	112.87	female	HET:PLS3hom+ASO	p10	
p10-167	AHSG	158.53	female	HET:PLS3hom+ASO	p10	
p10-244	AHSG	161.2	female	HET:PLS3hom+ASO	p10	
p10-121	AHSG	115.48	male	HET:PLS3hom+ASO	p10	
p10-126	AHSG	120.6	male	HET:PLS3hom+ASO	p10	
p10-166	AHSG	139.68	male	HET:PLS3hom+ASO	p10	
p10-201	AHSG	157.67	female	HET+ASO	p10	
p10-207	AHSG	175.19	female	HET+ASO	p10	
p10-208	AHSG	195.98	female	HET+ASO	p10	
p10-134	AHSG	150.93	male	HET+ASO	p10	
p10-135	AHSG	217.44	male	HET+ASO	p10	
p10-136	AHSG	216.5	male	HET+ASO	p10	

p10-164	AHSG	451.98	female	SMA	p10	
p10-182	AHSG	301.77	female	SMA	p10	
p10-193/3	AHSG	362	female	SMA	p10	
p10-131	AHSG	191.62	male	SMA	p10	
p10-132	AHSG	258.23	male	SMA	p10	
p10-133	AHSG	297.48	male	SMA	p10	
p10-144	AHSG	242.92	female	SMA:PLS3het	p10	
p10-146	AHSG	101.93	female	SMA:PLS3het	p10	
p10-151	AHSG	373.7	female	SMA:PLS3het	p10	
p10-143	AHSG	244.69	male	SMA:PLS3het	p10	
p10-145	AHSG	127.01	male	SMA:PLS3het	p10	
p10-161	AHSG	206.02	male	SMA:PLS3het	p10	
p10-127	AHSG	335.78	female	SMA:PLS3het+ASO	p10	
p10-155	AHSG	212.77	female	SMA:PLS3het+ASO	p10	
p10-220	AHSG	401.68	female	SMA:PLS3het+ASO	p10	
p10-154	AHSG	274.59	male	SMA:PLS3het+ASO	p10	
p10-156	AHSG	242.73	male	SMA:PLS3het+ASO	p10	
p10-219	AHSG	433.8	male	SMA:PLS3het+ASO	p10	
p10-147	AHSG	283.23	female	SMA:PLS3hom	p10	*
p10-153	AHSG	190.43	female	SMA:PLS3hom	p10	
p10-196	AHSG	176.39	female	SMA:PLS3hom	p10	
p10-176	AHSG	199.23	male	SMA:PLS3hom	p10	
p10-200	AHSG	216.69	male	SMA:PLS3hom	p10	
p10-226	AHSG	200.5	male	SMA:PLS3hom	p10	
p10-119	AHSG	277.99	female	SMA:PLS3hom+ASO	p10	
p10-218	AHSG	254.95	female	SMA:PLS3hom+ASO	p10	
p10-243	AHSG	239.16	female	SMA:PLS3hom+ASO	p10	
p10-118	AHSG	180.33	male	SMA:PLS3hom+ASO	p10	
p10-124	AHSG	196.2	male	SMA:PLS3hom+ASO	p10	
p10-125	AHSG	475.59	male	SMA:PLS3hom+ASO	p10	*
p10-138	AHSG	418.62	female	SMA+ASO	p10	
p10-140	AHSG	427.4	female	SMA+ASO	p10	
p10-209	AHSG	465.35	female	SMA+ASO	p10	
p10-137	AHSG	434.53	male	SMA+ASO	p10	
p10-139	AHSG	348.78	male	SMA+ASO	p10	
p10-210	AHSG	292.55	male	SMA+ASO	p10	
p10-202	AHSG	207.98	female	WT	p10	
p10-203	AHSG	173.23	female	WT	p10	
p10-204	AHSG	174.27	female	WT	p10	
p10-205	AHSG	192.19	male	WT	p10	
p10-206	AHSG	193.53	male	WT	p10	
p10-218	AHSG	215.66	male	WT	p10	
p10-212	AHSG	154.76	female	WT+ASO	p10	
p10-213	AHSG	157.42	female	WT+ASO	p10	
p10-214	AHSG	194.21	female	WT+ASO	p10	
p10-215	AHSG	157.91	male	WT+ASO	p10	

p10-216	AHSG	197.22	male	WT+ASO	p10	
p10-217	AHSG	209.83	male	WT+ASO	p10	
p21-150	AHSG	57.12	female	HET:PLS3het+ASO	p21	*
p21-163	AHSG	89.09	female	HET:PLS3het+ASO	p21	*
p21-178	AHSG	75.98	female	HET:PLS3het+ASO	p21	
p21-151	AHSG	74.14	male	HET:PLS3het+ASO	p21	
p21-157	AHSG	77.38	male	HET:PLS3het+ASO	p21	
p21-179	AHSG	76.18	male	HET:PLS3het+ASO	p21	
p21-167	AHSG	70.05	female	HET:PLS3hom+ASO	p21	
p21-211	AHSG	82.29	female	HET:PLS3hom+ASO	p21	
p21-223	AHSG	191.16	female	HET:PLS3hom+ASO	p21	*
p21-158	AHSG	91.01	male	HET:PLS3hom+ASO	p21	
p21-166	AHSG	81.76	male	HET:PLS3hom+ASO	p21	
p21-190	AHSG	66.3	male	HET:PLS3hom+ASO	p21	
p21-152	AHSG	69.99	female	HET+ASO	p21	
p21-153	AHSG	52.49	female	HET+ASO	p21	
p21-175	AHSG	61.07	female	HET+ASO	p21	
p21-149	AHSG	74.34	male	HET+ASO	p21	
p21-154	AHSG	71.93	male	HET+ASO	p21	
p21-159	AHSG	95.67	male	HET+ASO	p21	*
p21-156	AHSG	75.91	female	SMA:PLS3het+ASO	p21	
p21-164	AHSG	79.9	female	SMA:PLS3het+ASO	p21	
p21-168	AHSG	123.96	female	SMA:PLS3het+ASO	p21	
p21-155	AHSG	158.16	male	SMA:PLS3het+ASO	p21	
p21-165	AHSG	92.86	male	SMA:PLS3het+ASO	p21	
p21-169	AHSG	95.08	male	SMA:PLS3het+ASO	p21	
p21-196	AHSG	115.68	female	SMA:PLS3hom+ASO	p21	
p21-221	AHSG	122.46	female	SMA:PLS3hom+ASO	p21	
p21-222	AHSG	265.76	female	SMA:PLS3hom+ASO	p21	*
p21-160	AHSG	97.59	male	SMA:PLS3hom+ASO	p21	
p21-162	AHSG	123.52	male	SMA:PLS3hom+ASO	p21	
p21-212	AHSG	93.67	male	SMA:PLS3hom+ASO	p21	
p21-182	AHSG	147.08	female	SMA+ASO	p21	
p21-183	AHSG	62.53	female	SMA+ASO	p21	
p21-189	AHSG	140.66	female	SMA+ASO	p21	
p21-184	AHSG	98.86	male	SMA+ASO	p21	
p21-185	AHSG	163.86	male	SMA+ASO	p21	
p21-191	AHSG	57.71	male	SMA+ASO	p21	
p21-207	AHSG	69.08	female	WT+ASO	p21	
p21-214	AHSG	87.97	female	WT+ASO	p21	
p21-215	AHSG	77.96	female	WT+ASO	p21	
p21-197	AHSG	64.33	male	WT+ASO	p21	
p21-198	AHSG	81.7	male	WT+ASO	p21	
p21-208	AHSG	77.84	male	WT+ASO	p21	
p10-173	CLEC3B	14.33	female	HET	p10	
p10-174	CLEC3B	14.22	female	HET	p10	

p10-187	CLEC3B	14.29	female	HET	p10	
p10-128	CLEC3B	15.12	male	HET	p10	
p10-129	CLEC3B	14.18	male	HET	p10	
p10-130	CLEC3B	15.32	male	HET	p10	
p10-141	CLEC3B	15.72	female	HET:PLS3het	p10	
p10-142	CLEC3B	16.37	female	HET:PLS3het	p10	
p10-159	CLEC3B	16.75	female	HET:PLS3het	p10	
p10-149	CLEC3B	17.39	male	HET:PLS3het	p10	
p10-160	CLEC3B	14.57	male	HET:PLS3het	p10	
p10-172	CLEC3B	17.44	male	HET:PLS3het	p10	
p10-120	CLEC3B	11.44	female	HET:PLS3het+ASO	p10	
p10-157	CLEC3B	14.39	female	HET:PLS3het+ASO	p10	
p10-224	CLEC3B	16.86	female	HET:PLS3het+ASO	p10	
p10-123	CLEC3B	12.35	male	HET:PLS3het+ASO	p10	
p10-158	CLEC3B	13.22	male	HET:PLS3het+ASO	p10	
p10-223	CLEC3B	22.57	male	HET:PLS3het+ASO	p10	*
p10-148	CLEC3B	16	female	HET:PLS3hom	p10	
p10-181	CLEC3B	16.51	female	HET:PLS3hom	p10	
p10-194	CLEC3B	13.13	female	HET:PLS3hom	p10	*
p10-150	CLEC3B	16.35	male	HET:PLS3hom	p10	
p10-152	CLEC3B	16.67	male	HET:PLS3hom	p10	
p10-178	CLEC3B	15.8	male	HET:PLS3hom	p10	
p10-122	CLEC3B	13.48	female	HET:PLS3hom+ASO	p10	
p10-167	CLEC3B	15.62	female	HET:PLS3hom+ASO	p10	
p10-244	CLEC3B	16.1	female	HET:PLS3hom+ASO	p10	
p10-121	CLEC3B	12.6	male	HET:PLS3hom+ASO	p10	
p10-126	CLEC3B	14.72	male	HET:PLS3hom+ASO	p10	
p10-166	CLEC3B	15.44	male	HET:PLS3hom+ASO	p10	
p10-201	CLEC3B	12.86	female	HET+ASO	p10	
p10-207	CLEC3B	12.57	female	HET+ASO	p10	
p10-208	CLEC3B	12.64	female	HET+ASO	p10	
p10-134	CLEC3B	12.29	male	HET+ASO	p10	
p10-135	CLEC3B	17.2	male	HET+ASO	p10	
p10-136	CLEC3B	17.4	male	HET+ASO	p10	
p10-164	CLEC3B	16.46	female	SMA	p10	
p10-182	CLEC3B	14.92	female	SMA	p10	
p10-193/3	CLEC3B	17.76	female	SMA	p10	
p10-131	CLEC3B	11.47	male	SMA	p10	
p10-132	CLEC3B	14.79	male	SMA	p10	
p10-133	CLEC3B	13.14	male	SMA	p10	
p10-144	CLEC3B	16.72	female	SMA:PLS3het	p10	
p10-146	CLEC3B	15.72	female	SMA:PLS3het	p10	
p10-151	CLEC3B	13.51	female	SMA:PLS3het	p10	
p10-143	CLEC3B	13.38	male	SMA:PLS3het	p10	
p10-145	CLEC3B	17.11	male	SMA:PLS3het	p10	
p10-161	CLEC3B	14.02	male	SMA:PLS3het	p10	

p10-127	CLEC3B	9.83	female	SMA:PLS3het+ASO	p10	
p10-155	CLEC3B	11.27	female	SMA:PLS3het+ASO	p10	
p10-220	CLEC3B	11.37	female	SMA:PLS3het+ASO	p10	
p10-154	CLEC3B	9.73	male	SMA:PLS3het+ASO	p10	
p10-156	CLEC3B	10.92	male	SMA:PLS3het+ASO	p10	
p10-219	CLEC3B	10.24	male	SMA:PLS3het+ASO	p10	
p10-147	CLEC3B	13.42	female	SMA:PLS3hom	p10	
p10-153	CLEC3B	15.35	female	SMA:PLS3hom	p10	
p10-196	CLEC3B	11.02	female	SMA:PLS3hom	p10	
p10-176	CLEC3B	12.01	male	SMA:PLS3hom	p10	
p10-200	CLEC3B	13.15	male	SMA:PLS3hom	p10	
p10-226	CLEC3B	14.15	male	SMA:PLS3hom	p10	
p10-119	CLEC3B	5.95	female	SMA:PLS3hom+ASO	p10	
p10-218	CLEC3B	10.84	female	SMA:PLS3hom+ASO	p10	
p10-243	CLEC3B	12.08	female	SMA:PLS3hom+ASO	p10	
p10-118	CLEC3B	6.85	male	SMA:PLS3hom+ASO	p10	
p10-124	CLEC3B	8.98	male	SMA:PLS3hom+ASO	p10	
p10-125	CLEC3B	9.09	male	SMA:PLS3hom+ASO	p10	
p10-138	CLEC3B	11.25	female	SMA+ASO	p10	
p10-140	CLEC3B	10.4	female	SMA+ASO	p10	
p10-209	CLEC3B	10.15	female	SMA+ASO	p10	
p10-137	CLEC3B	12.6	male	SMA+ASO	p10	
p10-139	CLEC3B	14.85	male	SMA+ASO	p10	
p10-210	CLEC3B	7.63	male	SMA+ASO	p10	
p10-202	CLEC3B	18.77	female	WT	p10	
p10-203	CLEC3B	16.21	female	WT	p10	
p10-204	CLEC3B	9.88	female	WT	p10	
p10-205	CLEC3B	8.72	male	WT	p10	
p10-206	CLEC3B	9.25	male	WT	p10	
p10-218	CLEC3B	9.03	male	WT	p10	
p10-212	CLEC3B	19.27	female	WT+ASO	p10	
p10-213	CLEC3B	16.23	female	WT+ASO	p10	
p10-214	CLEC3B	18.85	female	WT+ASO	p10	
p10-215	CLEC3B	16.26	male	WT+ASO	p10	
p10-216	CLEC3B	18.99	male	WT+ASO	p10	
p10-217	CLEC3B	19.66	male	WT+ASO	p10	
p21-150	CLEC3B	12.85	female	HET:PLS3het+ASO	p21	
p21-163	CLEC3B	13.75	female	HET:PLS3het+ASO	p21	
p21-178	CLEC3B	10.73	female	HET:PLS3het+ASO	p21	
p21-151	CLEC3B	13.07	male	HET:PLS3het+ASO	p21	
p21-157	CLEC3B	9.73	male	HET:PLS3het+ASO	p21	
p21-179	CLEC3B	8.3	male	HET:PLS3het+ASO	p21	
p21-167	CLEC3B	10.16	female	HET:PLS3hom+ASO	p21	
p21-211	CLEC3B	9.49	female	HET:PLS3hom+ASO	p21	
p21-223	CLEC3B	9.34	female	HET:PLS3hom+ASO	p21	
p21-158	CLEC3B	11.74	male	HET:PLS3hom+ASO	p21	

p21-166	CLEC3B	12.43	male	HET:PLS3hom+ASO	p21	
p21-190	CLEC3B	9.06	male	HET:PLS3hom+ASO	p21	
p21-152	CLEC3B	17.71	female	HET+ASO	p21	*
p21-153	CLEC3B	11.21	female	HET+ASO	p21	
p21-175	CLEC3B	10.73	female	HET+ASO	p21	
p21-149	CLEC3B	11.92	male	HET+ASO	p21	
p21-154	CLEC3B	13.22	male	HET+ASO	p21	
p21-159	CLEC3B	12.65	male	HET+ASO	p21	
p21-156	CLEC3B	9.09	female	SMA:PLS3het+ASO	p21	
p21-164	CLEC3B	11.69	female	SMA:PLS3het+ASO	p21	
p21-168	CLEC3B	11.71	female	SMA:PLS3het+ASO	p21	
p21-155	CLEC3B	15.88	male	SMA:PLS3het+ASO	p21	*
p21-165	CLEC3B	8.47	male	SMA:PLS3het+ASO	p21	
p21-169	CLEC3B	11.47	male	SMA:PLS3het+ASO	p21	
p21-196	CLEC3B	13.29	female	SMA:PLS3hom+ASO	p21	
p21-221	CLEC3B	11.64	female	SMA:PLS3hom+ASO	p21	
p21-222	CLEC3B	10.05	female	SMA:PLS3hom+ASO	p21	
p21-160	CLEC3B	8.26	male	SMA:PLS3hom+ASO	p21	
p21-162	CLEC3B	11.85	male	SMA:PLS3hom+ASO	p21	
p21-212	CLEC3B	13.36	male	SMA:PLS3hom+ASO	p21	
p21-182	CLEC3B	12.26	female	SMA+ASO	p21	
p21-183	CLEC3B	10.65	female	SMA+ASO	p21	
p21-189	CLEC3B	11.53	female	SMA+ASO	p21	
p21-184	CLEC3B	12.23	male	SMA+ASO	p21	
p21-185	CLEC3B	8.65	male	SMA+ASO	p21	
p21-191	CLEC3B	9.54	male	SMA+ASO	p21	
p21-207	CLEC3B	0.08	female	WT+ASO	p21	*
p21-214	CLEC3B	15.55	female	WT+ASO	p21	
p21-215	CLEC3B	13.7	female	WT+ASO	p21	
p21-197	CLEC3B	12.32	male	WT+ASO	p21	
p21-198	CLEC3B	15.46	male	WT+ASO	p21	
p21-208	CLEC3B	12.35	male	WT+ASO	p21	
p10-173	COMP	311.73	female	HET	p10	
p10-174	COMP	347.34	female	HET	p10	
p10-187	COMP	331.27	female	HET	p10	
p10-128	COMP	348.63	male	HET	p10	
p10-129	COMP	335.64	male	HET	p10	
p10-130	COMP	331.01	male	HET	p10	
p10-141	COMP	370.88	female	HET:PLS3het	p10	
p10-142	COMP	314.97	female	HET:PLS3het	p10	
p10-159	COMP	358.08	female	HET:PLS3het	p10	
p10-149	COMP	369.78	male	HET:PLS3het	p10	
p10-160	COMP	322.38	male	HET:PLS3het	p10	
p10-172	COMP	338.95	male	HET:PLS3het	p10	
p10-120	COMP	235.22	female	HET:PLS3het+ASO	p10	
p10-157	COMP	312.18	female	HET:PLS3het+ASO	p10	

p10-224	COMP	337.06	female	HET:PLS3het+ASO	p10	
p10-123	COMP	303.54	male	HET:PLS3het+ASO	p10	
p10-158	COMP	302.78	male	HET:PLS3het+ASO	p10	
p10-223	COMP	261.83	male	HET:PLS3het+ASO	p10	
p10-148	COMP	364.33	female	HET:PLS3hom	p10	
p10-181	COMP	367.6	female	HET:PLS3hom	p10	
p10-194	COMP	349.8	female	HET:PLS3hom	p10	
p10-150	COMP	351.47	male	HET:PLS3hom	p10	
p10-152	COMP	349.02	male	HET:PLS3hom	p10	
p10-178	COMP	327.98	male	HET:PLS3hom	p10	*
p10-122	COMP	293.95	female	HET:PLS3hom+ASO	p10	
p10-167	COMP	326.61	female	HET:PLS3hom+ASO	p10	
p10-244	COMP	342.58	female	HET:PLS3hom+ASO	p10	
p10-121	COMP	263.59	male	HET:PLS3hom+ASO	p10	
p10-126	COMP	294.34	male	HET:PLS3hom+ASO	p10	
p10-166	COMP	282.73	male	HET:PLS3hom+ASO	p10	
p10-201	COMP	287.25	female	HET+ASO	p10	
p10-207	COMP	281.3	female	HET+ASO	p10	
p10-208	COMP	290.21	female	HET+ASO	p10	
p10-134	COMP	303.09	male	HET+ASO	p10	
p10-135	COMP	322.96	male	HET+ASO	p10	
p10-136	COMP	345.85	male	HET+ASO	p10	
p10-164	COMP	226.74	female	SMA	p10	
p10-182	COMP	132.44	female	SMA	p10	
p10-193/3	COMP	147.26	female	SMA	p10	
p10-131	COMP	124.01	male	SMA	p10	
p10-132	COMP	180.84	male	SMA	p10	
p10-133	COMP	134.7	male	SMA	p10	
p10-144	COMP	115.96	female	SMA:PLS3het	p10	
p10-146	COMP	225.76	female	SMA:PLS3het	p10	
p10-151	COMP	201.54	female	SMA:PLS3het	p10	
p10-143	COMP	134.66	male	SMA:PLS3het	p10	
p10-145	COMP	33.7	male	SMA:PLS3het	p10	
p10-161	COMP	125.67	male	SMA:PLS3het	p10	
p10-127	COMP	136.12	female	SMA:PLS3het+ASO	p10	
p10-155	COMP	158.86	female	SMA:PLS3het+ASO	p10	
p10-220	COMP	143.27	female	SMA:PLS3het+ASO	p10	
p10-154	COMP	127.04	male	SMA:PLS3het+ASO	p10	
p10-156	COMP	140.57	male	SMA:PLS3het+ASO	p10	
p10-219	COMP	110.04	male	SMA:PLS3het+ASO	p10	
p10-147	COMP	136.49	female	SMA:PLS3hom	p10	
p10-153	COMP	41.18	female	SMA:PLS3hom	p10	
p10-196	COMP	61.46	female	SMA:PLS3hom	p10	
p10-176	COMP	64.27	male	SMA:PLS3hom	p10	
p10-200	COMP	79.91	male	SMA:PLS3hom	p10	
p10-226	COMP	279.36	male	SMA:PLS3hom	p10	*

p10-119	COMP	150.44	female	SMA:PLS3hom+ASO	p10	
p10-218	COMP	151.82	female	SMA:PLS3hom+ASO	p10	
p10-243	COMP	100.12	female	SMA:PLS3hom+ASO	p10	
p10-118	COMP	112.76	male	SMA:PLS3hom+ASO	p10	
p10-124	COMP	25.39	male	SMA:PLS3hom+ASO	p10	*
p10-125	COMP	172.19	male	SMA:PLS3hom+ASO	p10	
p10-138	COMP	150.84	female	SMA+ASO	p10	
p10-140	COMP	124.16	female	SMA+ASO	p10	
p10-209	COMP	100.36	female	SMA+ASO	p10	
p10-137	COMP	166.94	male	SMA+ASO	p10	
p10-139	COMP	155.71	male	SMA+ASO	p10	
p10-210	COMP	46.97	male	SMA+ASO	p10	
p10-202	COMP	389.39	female	WT	p10	
p10-203	COMP	329.8	female	WT	p10	
p10-204	COMP	274.72	female	WT	p10	
p10-205	COMP	277.53	male	WT	p10	
p10-206	COMP	245.47	male	WT	p10	
p10-218	COMP	255.7	male	WT	p10	
p10-212	COMP	406.21	female	WT+ASO	p10	
p10-213	COMP	299.58	female	WT+ASO	p10	
p10-214	COMP	370.05	female	WT+ASO	p10	
p10-215	COMP	322.78	male	WT+ASO	p10	
p10-216	COMP	384.97	male	WT+ASO	p10	
p10-217	COMP	386.23	male	WT+ASO	p10	
p21-150	COMP	162.5	female	HET:PLS3het+ASO	p21	
p21-163	COMP	214.12	female	HET:PLS3het+ASO	p21	*
p21-178	COMP	132.8	female	HET:PLS3het+ASO	p21	*
p21-151	COMP	167.18	male	HET:PLS3het+ASO	p21	
p21-157	COMP	160.54	male	HET:PLS3het+ASO	p21	
p21-179	COMP	155.16	male	HET:PLS3het+ASO	p21	
p21-167	COMP	171.45	female	HET:PLS3hom+ASO	p21	
p21-211	COMP	168.39	female	HET:PLS3hom+ASO	p21	
p21-223	COMP	243.02	female	HET:PLS3hom+ASO	p21	*
p21-158	COMP	179.19	male	HET:PLS3hom+ASO	p21	
p21-166	COMP	159.45	male	HET:PLS3hom+ASO	p21	
p21-190	COMP	107.19	male	HET:PLS3hom+ASO	p21	*
p21-152	COMP	246.83	female	HET+ASO	p21	*
p21-153	COMP	141.3	female	HET+ASO	p21	
p21-175	COMP	140.47	female	HET+ASO	p21	
p21-149	COMP	173.48	male	HET+ASO	p21	
p21-154	COMP	161.97	male	HET+ASO	p21	
p21-159	COMP	178.34	male	HET+ASO	p21	
p21-156	COMP	43.41	female	SMA:PLS3het+ASO	p21	
p21-164	COMP	122.03	female	SMA:PLS3het+ASO	p21	
p21-168	COMP	113.57	female	SMA:PLS3het+ASO	p21	
p21-155	COMP	66.86	male	SMA:PLS3het+ASO	p21	

p21-165	COMP	89.59	male	SMA:PLS3het+ASO	p21	
p21-169	COMP	111.7	male	SMA:PLS3het+ASO	p21	
p21-196	COMP	179.71	female	SMA:PLS3hom+ASO	p21	
p21-221	COMP	131.18	female	SMA:PLS3hom+ASO	p21	
p21-222	COMP	87.78	female	SMA:PLS3hom+ASO	p21	
p21-160	COMP	70.26	male	SMA:PLS3hom+ASO	p21	
p21-162	COMP	138.31	male	SMA:PLS3hom+ASO	p21	
p21-212	COMP	145.62	male	SMA:PLS3hom+ASO	p21	
p21-182	COMP	118.85	female	SMA+ASO	p21	
p21-183	COMP	94.47	female	SMA+ASO	p21	
p21-189	COMP	137.36	female	SMA+ASO	p21	
p21-184	COMP	136.38	male	SMA+ASO	p21	
p21-185	COMP	47.95	male	SMA+ASO	p21	
p21-191	COMP	48.45	male	SMA+ASO	p21	
p21-207	COMP	175.81	female	WT+ASO	p21	
p21-214	COMP	174.56	female	WT+ASO	p21	
p21-215	COMP	107.59	female	WT+ASO	p21	*
p21-197	COMP	154.29	male	WT+ASO	p21	
p21-198	COMP	181.13	male	WT+ASO	p21	
p21-208	COMP	151.83	male	WT+ASO	p21	
p10-173	DPP4	79.33	female	HET	p10	
p10-174	DPP4	82.06	female	HET	p10	
p10-187	DPP4	75.25	female	HET	p10	
p10-128	DPP4	84.97	male	HET	p10	
p10-129	DPP4	78.1	male	HET	p10	
p10-130	DPP4	81.08	male	HET	p10	
p10-141	DPP4	82.25	female	HET:PLS3het	p10	
p10-142	DPP4	78.22	female	HET:PLS3het	p10	
p10-159	DPP4	94.34	female	HET:PLS3het	p10	
p10-149	DPP4	86.65	male	HET:PLS3het	p10	
p10-160	DPP4	80.44	male	HET:PLS3het	p10	
p10-172	DPP4	95.43	male	HET:PLS3het	p10	
p10-120	DPP4	26.26	female	HET:PLS3het+ASO	p10	
p10-157	DPP4	33.71	female	HET:PLS3het+ASO	p10	
p10-224	DPP4	79.28	female	HET:PLS3het+ASO	p10	
p10-123	DPP4	41.39	male	HET:PLS3het+ASO	p10	
p10-158	DPP4	32.16	male	HET:PLS3het+ASO	p10	
p10-223	DPP4	92.26	male	HET:PLS3het+ASO	p10	
p10-148	DPP4	78.13	female	HET:PLS3hom	p10	
p10-181	DPP4	77.78	female	HET:PLS3hom	p10	
p10-194	DPP4	76.78	female	HET:PLS3hom	p10	
p10-150	DPP4	80.65	male	HET:PLS3hom	p10	
p10-152	DPP4	81.63	male	HET:PLS3hom	p10	
p10-178	DPP4	76.05	male	HET:PLS3hom	p10	
p10-122	DPP4	32.36	female	HET:PLS3hom+ASO	p10	
p10-167	DPP4	50.92	female	HET:PLS3hom+ASO	p10	

p10-244	DPP4	41.74	female	HET:PLS3hom+ASO	p10	
p10-121	DPP4	31.61	male	HET:PLS3hom+ASO	p10	
p10-126	DPP4	34.01	male	HET:PLS3hom+ASO	p10	
p10-166	DPP4	45.84	male	HET:PLS3hom+ASO	p10	
p10-201	DPP4	76.13	female	HET+ASO	p10	*
p10-207	DPP4	47.87	female	HET+ASO	p10	
p10-208	DPP4	42.79	female	HET+ASO	p10	
p10-134	DPP4	39.91	male	HET+ASO	p10	
p10-135	DPP4	40.47	male	HET+ASO	p10	
p10-136	DPP4	46.01	male	HET+ASO	p10	
p10-164	DPP4	86.19	female	SMA	p10	
p10-182	DPP4	65.09	female	SMA	p10	
p10-193/3	DPP4	128.08	female	SMA	p10	*
p10-131	DPP4	54.31	male	SMA	p10	
p10-132	DPP4	67.81	male	SMA	p10	
p10-133	DPP4	61.04	male	SMA	p10	
p10-144	DPP4	21.42	female	SMA:PLS3het	p10	
p10-146	DPP4	27.76	female	SMA:PLS3het	p10	
p10-151	DPP4	79.98	female	SMA:PLS3het	p10	
p10-143	DPP4	67.29	male	SMA:PLS3het	p10	
p10-145	DPP4	53.52	male	SMA:PLS3het	p10	
p10-161	DPP4	59.69	male	SMA:PLS3het	p10	
p10-127	DPP4	22.51	female	SMA:PLS3het+ASO	p10	
p10-155	DPP4	30.12	female	SMA:PLS3het+ASO	p10	
p10-220	DPP4	34.77	female	SMA:PLS3het+ASO	p10	
p10-154	DPP4	28.01	male	SMA:PLS3het+ASO	p10	
p10-156	DPP4	24.31	male	SMA:PLS3het+ASO	p10	
p10-219	DPP4	29.76	male	SMA:PLS3het+ASO	p10	
p10-147	DPP4	34.2	female	SMA:PLS3hom	p10	
p10-153	DPP4	16.57	female	SMA:PLS3hom	p10	
p10-196	DPP4	52.21	female	SMA:PLS3hom	p10	
p10-176	DPP4	45.67	male	SMA:PLS3hom	p10	
p10-200	DPP4	53.6	male	SMA:PLS3hom	p10	
p10-226	DPP4	49.98	male	SMA:PLS3hom	p10	
p10-119	DPP4	61.47	female	SMA:PLS3hom+ASO	p10	
p10-218	DPP4	83.09	female	SMA:PLS3hom+ASO	p10	
p10-243	DPP4	63.62	female	SMA:PLS3hom+ASO	p10	
p10-118	DPP4	56.3	male	SMA:PLS3hom+ASO	p10	
p10-124	DPP4	13.74	male	SMA:PLS3hom+ASO	p10	*
p10-125	DPP4	39.3	male	SMA:PLS3hom+ASO	p10	
p10-138	DPP4	35.04	female	SMA+ASO	p10	
p10-140	DPP4	33.13	female	SMA+ASO	p10	
p10-209	DPP4	36.21	female	SMA+ASO	p10	
p10-137	DPP4	32.55	male	SMA+ASO	p10	
p10-139	DPP4	52.82	male	SMA+ASO	p10	*
p10-210	DPP4	21.54	male	SMA+ASO	p10	*

p10-202	DPP4	104.8	female	WT	p10	
p10-203	DPP4	93.05	female	WT	p10	
p10-204	DPP4	85.51	female	WT	p10	
p10-205	DPP4	117.04	male	WT	p10	
p10-206	DPP4	96.23	male	WT	p10	
p10-218	DPP4	112.23	male	WT	p10	
p10-212	DPP4	86.77	female	WT+ASO	p10	
p10-213	DPP4	63.52	female	WT+ASO	p10	
p10-214	DPP4	75.3	female	WT+ASO	p10	
p10-215	DPP4	102.64	male	WT+ASO	p10	
p10-216	DPP4	106.6	male	WT+ASO	p10	
p10-217	DPP4	77.51	male	WT+ASO	p10	
p21-150	DPP4	76.26	female	HET:PLS3het+ASO	p21	
p21-163	DPP4	95.46	female	HET:PLS3het+ASO	p21	
p21-178	DPP4	78.15	female	HET:PLS3het+ASO	p21	
p21-151	DPP4	86.64	male	HET:PLS3het+ASO	p21	
p21-157	DPP4	85.67	male	HET:PLS3het+ASO	p21	
p21-179	DPP4	97.31	male	HET:PLS3het+ASO	p21	
p21-167	DPP4	99.07	female	HET:PLS3hom+ASO	p21	
p21-211	DPP4	113.15	female	HET:PLS3hom+ASO	p21	
p21-223	DPP4	69.32	female	HET:PLS3hom+ASO	p21	
p21-158	DPP4	101.48	male	HET:PLS3hom+ASO	p21	
p21-166	DPP4	69.39	male	HET:PLS3hom+ASO	p21	
p21-190	DPP4	84.34	male	HET:PLS3hom+ASO	p21	
p21-152	DPP4	160.43	female	HET+ASO	p21	*
p21-153	DPP4	74.37	female	HET+ASO	p21	
p21-175	DPP4	113.44	female	HET+ASO	p21	
p21-149	DPP4	84.97	male	HET+ASO	p21	
p21-154	DPP4	80.35	male	HET+ASO	p21	
p21-159	DPP4	93.74	male	HET+ASO	p21	
p21-156	DPP4	36.94	female	SMA:PLS3het+ASO	p21	
p21-164	DPP4	60.32	female	SMA:PLS3het+ASO	p21	
p21-168	DPP4	56.82	female	SMA:PLS3het+ASO	p21	
p21-155	DPP4	86.05	male	SMA:PLS3het+ASO	p21	*
p21-165	DPP4	39.34	male	SMA:PLS3het+ASO	p21	
p21-169	DPP4	60.59	male	SMA:PLS3het+ASO	p21	
p21-196	DPP4	71.81	female	SMA:PLS3hom+ASO	p21	
p21-221	DPP4	82	female	SMA:PLS3hom+ASO	p21	
p21-222	DPP4	54.86	female	SMA:PLS3hom+ASO	p21	
p21-160	DPP4	49.72	male	SMA:PLS3hom+ASO	p21	
p21-162	DPP4	75	male	SMA:PLS3hom+ASO	p21	
p21-212	DPP4	72.2	male	SMA:PLS3hom+ASO	p21	
p21-182	DPP4	59.8	female	SMA+ASO	p21	
p21-183	DPP4	61.46	female	SMA+ASO	p21	
p21-189	DPP4	91.47	female	SMA+ASO	p21	*
p21-184	DPP4	67.73	male	SMA+ASO	p21	

p21-185	DPP4	26.03	male	SMA+ASO	p21	*
p21-191	DPP4	62.85	male	SMA+ASO	p21	
p21-207	DPP4	129	female	WT+ASO	p21	
p21-214	DPP4	208.24	female	WT+ASO	p21	
p21-215	DPP4	157.65	female	WT+ASO	p21	
p21-197	DPP4	86.97	male	WT+ASO	p21	
p21-198	DPP4	106.69	male	WT+ASO	p21	
p21-208	DPP4	109	male	WT+ASO	p21	
p10-173	SMN	42157	female	HET	p10	
p10-174	SMN	33626	female	HET	p10	
p10-187	SMN	37344	female	HET	p10	
p10-128	SMN	5766	male	HET	p10	
p10-129	SMN	24038	male	HET	p10	
p10-130	SMN	21501	male	HET	p10	
p10-141	SMN	22491	female	HET:PLS3het	p10	
p10-142	SMN	37071	female	HET:PLS3het	p10	
p10-159	SMN	16757	female	HET:PLS3het	p10	
p10-149	SMN	46327	male	HET:PLS3het	p10	
p10-160	SMN	13282	male	HET:PLS3het	p10	
p10-172	SMN	18125	male	HET:PLS3het	p10	
p10-120	SMN	19410	female	HET:PLS3het+ASO	p10	
p10-157	SMN	24929	female	HET:PLS3het+ASO	p10	
p10-224	SMN	14375	female	HET:PLS3het+ASO	p10	
p10-123	SMN	14039	male	HET:PLS3het+ASO	p10	
p10-158	SMN	23181	male	HET:PLS3het+ASO	p10	
p10-223	SMN	105954	male	HET:PLS3het+ASO	p10	*
p10-148	SMN	82404	female	HET:PLS3hom	p10	*
p10-181	SMN	36046	female	HET:PLS3hom	p10	
p10-194	SMN	27016	female	HET:PLS3hom	p10	
p10-150	SMN	38744	male	HET:PLS3hom	p10	
p10-152	SMN	52350	male	HET:PLS3hom	p10	
p10-178	SMN	53686	male	HET:PLS3hom	p10	
p10-122	SMN	11776	female	HET:PLS3hom+ASO	p10	
p10-167	SMN	67497	female	HET:PLS3hom+ASO	p10	
p10-244	SMN	80941	female	HET:PLS3hom+ASO	p10	
p10-121	SMN	16446	male	HET:PLS3hom+ASO	p10	
p10-126	SMN	10970	male	HET:PLS3hom+ASO	p10	
p10-166	SMN	80036	male	HET:PLS3hom+ASO	p10	
p10-201	SMN	14356	female	HET+ASO	p10	
p10-207	SMN	34965	female	HET+ASO	p10	
p10-208	SMN	21897	female	HET+ASO	p10	
p10-134	SMN	44059	male	HET+ASO	p10	
p10-135	SMN	61441	male	HET+ASO	p10	
p10-136	SMN	39686	male	HET+ASO	p10	
p10-164	SMN	1097	female	SMA	p10	*
p10-182	SMN	9467	female	SMA	p10	

p10-193/3	SMN	6375	female	SMA	p10	
p10-131	SMN	10880	male	SMA	p10	
p10-132	SMN	13575	male	SMA	p10	
p10-133	SMN	8866	male	SMA	p10	
p10-144	SMN	7951	female	SMA:PLS3het	p10	
p10-146	SMN	22657	female	SMA:PLS3het	p10	
p10-151	SMN	12200	female	SMA:PLS3het	p10	
p10-143	SMN	16652	male	SMA:PLS3het	p10	
p10-145	SMN	20407	male	SMA:PLS3het	p10	
p10-161	SMN	3467	male	SMA:PLS3het	p10	
p10-127	SMN	1492	female	SMA:PLS3het+ASO	p10	
p10-155	SMN	12267	female	SMA:PLS3het+ASO	p10	
p10-220	SMN	19185	female	SMA:PLS3het+ASO	p10	
p10-154	SMN	18044	male	SMA:PLS3het+ASO	p10	
p10-156	SMN	3150	male	SMA:PLS3het+ASO	p10	
p10-219	SMN	33075	male	SMA:PLS3het+ASO	p10	
p10-147	SMN	17775	female	SMA:PLS3hom	p10	
p10-153	SMN	14392	female	SMA:PLS3hom	p10	
p10-196	SMN	2207	female	SMA:PLS3hom	p10	
p10-176	SMN	9894	male	SMA:PLS3hom	p10	
p10-200	SMN	2363	male	SMA:PLS3hom	p10	
p10-226	SMN	12724	male	SMA:PLS3hom	p10	
p10-119	SMN	2964	female	SMA:PLS3hom+ASO	p10	
p10-218	SMN	19510	female	SMA:PLS3hom+ASO	p10	
p10-243	SMN	16316	female	SMA:PLS3hom+ASO	p10	
p10-118	SMN	898	male	SMA:PLS3hom+ASO	p10	
p10-124	SMN	1172	male	SMA:PLS3hom+ASO	p10	
p10-125	SMN	18572	male	SMA:PLS3hom+ASO	p10	
p10-138	SMN	13031	female	SMA+ASO	p10	
p10-140	SMN	16258	female	SMA+ASO	p10	
p10-209	SMN	9467	female	SMA+ASO	p10	
p10-137	SMN	10273	male	SMA+ASO	p10	
p10-139	SMN	2531	male	SMA+ASO	p10	
p10-210	SMN	6927	male	SMA+ASO	p10	
p10-202	SMN	32000	female	WT	p10	
p10-203	SMN	39993	female	WT	p10	
p10-204	SMN	56403	female	WT	p10	
p10-205	SMN	62980	male	WT	p10	
p10-206	SMN	67797	male	WT	p10	
p10-218	SMN	41436	male	WT	p10	
p10-212	SMN	49991	female	WT+ASO	p10	
p10-213	SMN	50545	female	WT+ASO	p10	
p10-214	SMN	48980	female	WT+ASO	p10	
p10-215	SMN	63156	male	WT+ASO	p10	
p10-216	SMN	57931	male	WT+ASO	p10	
p10-217	SMN	78452	male	WT+ASO	p10	

p21-150	SMN	30667	female	HET:PLS3het+ASO	p21	
p21-163	SMN	50029	female	HET:PLS3het+ASO	p21	*
p21-178	SMN	16247	female	HET:PLS3het+ASO	p21	
p21-151	SMN	32299	male	HET:PLS3het+ASO	p21	
p21-157	SMN	30590	male	HET:PLS3het+ASO	p21	
p21-179	SMN	21370	male	HET:PLS3het+ASO	p21	
p21-167	SMN	18752	female	HET:PLS3hom+ASO	p21	
p21-211	SMN	41725	female	HET:PLS3hom+ASO	p21	
p21-223	SMN	192246	female	HET:PLS3hom+ASO	p21	*
p21-158	SMN	22794	male	HET:PLS3hom+ASO	p21	
p21-166	SMN	21479	male	HET:PLS3hom+ASO	p21	
p21-190	SMN	15450	male	HET:PLS3hom+ASO	p21	
p21-152	SMN	8035	female	HET+ASO	p21	
p21-153	SMN	56288	female	HET+ASO	p21	
p21-175	SMN	44031	female	HET+ASO	p21	
p21-149	SMN	14233	male	HET+ASO	p21	
p21-154	SMN	40087	male	HET+ASO	p21	
p21-159	SMN	20921	male	HET+ASO	p21	
p21-156	SMN	5295	female	SMA:PLS3het+ASO	p21	
p21-164	SMN	5316	female	SMA:PLS3het+ASO	p21	
p21-168	SMN	6170	female	SMA:PLS3het+ASO	p21	
p21-155	SMN	1432	male	SMA:PLS3het+ASO	p21	*
p21-165	SMN	4598	male	SMA:PLS3het+ASO	p21	
p21-169	SMN	5551	male	SMA:PLS3het+ASO	p21	
p21-196	SMN	5023	female	SMA:PLS3hom+ASO	p21	
p21-221	SMN	16446	female	SMA:PLS3hom+ASO	p21	
p21-222	SMN	36014	female	SMA:PLS3hom+ASO	p21	*
p21-160	SMN	2742	male	SMA:PLS3hom+ASO	p21	
p21-162	SMN	9052	male	SMA:PLS3hom+ASO	p21	
p21-212	SMN	7670	male	SMA:PLS3hom+ASO	p21	
p21-182	SMN	11583	female	SMA+ASO	p21	
p21-183	SMN	5252	female	SMA+ASO	p21	
p21-189	SMN	9898	female	SMA+ASO	p21	
p21-184	SMN	17596	male	SMA+ASO	p21	*
p21-185	SMN	9779	male	SMA+ASO	p21	
p21-191	SMN	7027	male	SMA+ASO	p21	
p21-207	SMN	86455	female	WT+ASO	p21	
p21-214	SMN	33167	female	WT+ASO	p21	
p21-215	SMN	60930	female	WT+ASO	p21	
p21-197	SMN	163364	male	WT+ASO	p21	*
p21-198	SMN	74197	male	WT+ASO	p21	
p21-208	SMN	94138	male	WT+ASO	p21	
p10-173	SPP1	182.47	female	HET	p10	
p10-174	SPP1	198.56	female	HET	p10	
p10-187	SPP1	214.31	female	HET	p10	
p10-128	SPP1	241.33	male	HET	p10	*

p10-129	SPP1	204.03	male	HET	p10	
p10-130	SPP1	196.43	male	HET	p10	
p10-141	SPP1	272.63	female	HET:PLS3het	p10	
p10-142	SPP1	212.14	female	HET:PLS3het	p10	
p10-159	SPP1	209.42	female	HET:PLS3het	p10	
p10-149	SPP1	322.56	male	HET:PLS3het	p10	
p10-160	SPP1	192.7	male	HET:PLS3het	p10	
p10-172	SPP1	224.2	male	HET:PLS3het	p10	
p10-120	SPP1	218.24	female	HET:PLS3het+ASO	p10	
p10-157	SPP1	232.1	female	HET:PLS3het+ASO	p10	
p10-224	SPP1	327.39	female	HET:PLS3het+ASO	p10	
p10-123	SPP1	282.63	male	HET:PLS3het+ASO	p10	
p10-158	SPP1	214.56	male	HET:PLS3het+ASO	p10	
p10-223	SPP1	169.39	male	HET:PLS3het+ASO	p10	
p10-148	SPP1	293.43	female	HET:PLS3hom	p10	
p10-181	SPP1	272.94	female	HET:PLS3hom	p10	
p10-194	SPP1	215.9	female	HET:PLS3hom	p10	*
p10-150	SPP1	276.42	male	HET:PLS3hom	p10	
p10-152	SPP1	268.61	male	HET:PLS3hom	p10	
p10-178	SPP1	297.61	male	HET:PLS3hom	p10	
p10-122	SPP1	281.75	female	HET:PLS3hom+ASO	p10	
p10-167	SPP1	260.63	female	HET:PLS3hom+ASO	p10	
p10-244	SPP1	284.3	female	HET:PLS3hom+ASO	p10	
p10-121	SPP1	265.68	male	HET:PLS3hom+ASO	p10	
p10-126	SPP1	320.07	male	HET:PLS3hom+ASO	p10	*
p10-166	SPP1	285.79	male	HET:PLS3hom+ASO	p10	
p10-201	SPP1	151.78	female	HET+ASO	p10	
p10-207	SPP1	165.61	female	HET+ASO	p10	
p10-208	SPP1	178.45	female	HET+ASO	p10	
p10-134	SPP1	325.76	male	HET+ASO	p10	
p10-135	SPP1	209.72	male	HET+ASO	p10	
p10-136	SPP1	245.02	male	HET+ASO	p10	
p10-164	SPP1	251.07	female	SMA	p10	
p10-182	SPP1	205.65	female	SMA	p10	
p10-193/3	SPP1	613.3	female	SMA	p10	*
p10-131	SPP1	168.25	male	SMA	p10	
p10-132	SPP1	216.17	male	SMA	p10	
p10-133	SPP1	209.83	male	SMA	p10	
p10-144	SPP1	153.73	female	SMA:PLS3het	p10	
p10-146	SPP1	205.86	female	SMA:PLS3het	p10	
p10-151	SPP1	364.79	female	SMA:PLS3het	p10	
p10-143	SPP1	304.44	male	SMA:PLS3het	p10	
p10-145	SPP1	125.04	male	SMA:PLS3het	p10	
p10-161	SPP1	294.07	male	SMA:PLS3het	p10	
p10-127	SPP1	178.34	female	SMA:PLS3het+ASO	p10	
p10-155	SPP1	182.94	female	SMA:PLS3het+ASO	p10	

p10-220	SPP1	214.15	female	SMA:PLS3het+ASO	p10	
p10-154	SPP1	179.56	male	SMA:PLS3het+ASO	p10	
p10-156	SPP1	173.32	male	SMA:PLS3het+ASO	p10	
p10-219	SPP1	207.72	male	SMA:PLS3het+ASO	p10	
p10-147	SPP1	209.25	female	SMA:PLS3hom	p10	
p10-153	SPP1	87.67	female	SMA:PLS3hom	p10	*
p10-196	SPP1	163.26	female	SMA:PLS3hom	p10	
p10-176	SPP1	171.66	male	SMA:PLS3hom	p10	
p10-200	SPP1	211.61	male	SMA:PLS3hom	p10	
p10-226	SPP1	198.1	male	SMA:PLS3hom	p10	
p10-119	SPP1	268.54	female	SMA:PLS3hom+ASO	p10	
p10-218	SPP1	371.8	female	SMA:PLS3hom+ASO	p10	*
p10-243	SPP1	289.67	female	SMA:PLS3hom+ASO	p10	
p10-118	SPP1	251.11	male	SMA:PLS3hom+ASO	p10	
p10-124	SPP1	70.84	male	SMA:PLS3hom+ASO	p10	*
p10-125	SPP1	277	male	SMA:PLS3hom+ASO	p10	
p10-138	SPP1	183.56	female	SMA+ASO	p10	
p10-140	SPP1	170.88	female	SMA+ASO	p10	
p10-209	SPP1	212.88	female	SMA+ASO	p10	
p10-137	SPP1	209.91	male	SMA+ASO	p10	
p10-139	SPP1	206.29	male	SMA+ASO	p10	
p10-210	SPP1	124.73	male	SMA+ASO	p10	
p10-202	SPP1	304.09	female	WT	p10	
p10-203	SPP1	294.21	female	WT	p10	
p10-204	SPP1	342.26	female	WT	p10	
p10-205	SPP1	294.05	male	WT	p10	
p10-206	SPP1	391.57	male	WT	p10	*
p10-218	SPP1	268.4	male	WT	p10	
p10-212	SPP1	378.36	female	WT+ASO	p10	
p10-213	SPP1	263.47	female	WT+ASO	p10	
p10-214	SPP1	366	female	WT+ASO	p10	
p10-215	SPP1	314.92	male	WT+ASO	p10	
p10-216	SPP1	420.46	male	WT+ASO	p10	
p10-217	SPP1	437.9	male	WT+ASO	p10	
p21-150	SPP1	119.43	female	HET:PLS3het+ASO	p21	
p21-163	SPP1	287.51	female	HET:PLS3het+ASO	p21	*
p21-178	SPP1	125.85	female	HET:PLS3het+ASO	p21	
p21-151	SPP1	156.4	male	HET:PLS3het+ASO	p21	
p21-157	SPP1	171.52	male	HET:PLS3het+ASO	p21	
p21-179	SPP1	144.55	male	HET:PLS3het+ASO	p21	
p21-167	SPP1	200.16	female	HET:PLS3hom+ASO	p21	
p21-211	SPP1	205.45	female	HET:PLS3hom+ASO	p21	
p21-223	SPP1	270.07	female	HET:PLS3hom+ASO	p21	
p21-158	SPP1	234.15	male	HET:PLS3hom+ASO	p21	
p21-166	SPP1	243.44	male	HET:PLS3hom+ASO	p21	
p21-190	SPP1	111.11	male	HET:PLS3hom+ASO	p21	*

p21-152	SPP1	170.41	female	HET+ASO	p21	
p21-153	SPP1	114.36	female	HET+ASO	p21	
p21-175	SPP1	100.6	female	HET+ASO	p21	
p21-149	SPP1	166.4	male	HET+ASO	p21	
p21-154	SPP1	130.27	male	HET+ASO	p21	
p21-159	SPP1	189.43	male	HET+ASO	p21	
p21-156	SPP1	493.88	female	SMA:PLS3het+ASO	p21	
p21-164	SPP1	205.32	female	SMA:PLS3het+ASO	p21	
p21-168	SPP1	221.61	female	SMA:PLS3het+ASO	p21	
p21-155	SPP1	1024.01	male	SMA:PLS3het+ASO	p21	*
p21-165	SPP1	134.1	male	SMA:PLS3het+ASO	p21	
p21-169	SPP1	218.26	male	SMA:PLS3het+ASO	p21	
p21-196	SPP1	217.76	female	SMA:PLS3hom+ASO	p21	
p21-221	SPP1	236.51	female	SMA:PLS3hom+ASO	p21	
p21-222	SPP1	349.86	female	SMA:PLS3hom+ASO	p21	
p21-160	SPP1	262.65	male	SMA:PLS3hom+ASO	p21	
p21-162	SPP1	288.18	male	SMA:PLS3hom+ASO	p21	
p21-212	SPP1	298.73	male	SMA:PLS3hom+ASO	p21	
p21-182	SPP1	280.62	female	SMA+ASO	p21	
p21-183	SPP1	191.91	female	SMA+ASO	p21	
p21-189	SPP1	265.47	female	SMA+ASO	p21	
p21-184	SPP1	182.28	male	SMA+ASO	p21	
p21-185	SPP1	480.4	male	SMA+ASO	p21	*
p21-191	SPP1	160.56	male	SMA+ASO	p21	
p21-207	SPP1	191.94	female	WT+ASO	p21	
p21-214	SPP1	228.78	female	WT+ASO	p21	
p21-215	SPP1	208.95	female	WT+ASO	p21	
p21-197	SPP1	116.89	male	WT+ASO	p21	
p21-198	SPP1	148.54	male	WT+ASO	p21	
p21-208	SPP1	140.62	male	WT+ASO	p21	
p10-173	VTN	10998.53	female	HET	p10	
p10-174	VTN	12735.22	female	HET	p10	
p10-187	VTN	12293.84	female	HET	p10	
p10-128	VTN	12458.15	male	HET	p10	
p10-129	VTN	9929.1	male	HET	p10	
p10-130	VTN	11519.99	male	HET	p10	
p10-141	VTN	13155.17	female	HET:PLS3het	p10	
p10-142	VTN	12646.47	female	HET:PLS3het	p10	
p10-159	VTN	13460.36	female	HET:PLS3het	p10	
p10-149	VTN	13617.66	male	HET:PLS3het	p10	
p10-160	VTN	14040.96	male	HET:PLS3het	p10	
p10-172	VTN	12954.21	male	HET:PLS3het	p10	
p10-120	VTN	7907.43	female	HET:PLS3het+ASO	p10	
p10-157	VTN	11627.81	female	HET:PLS3het+ASO	p10	
p10-224	VTN	12048.92	female	HET:PLS3het+ASO	p10	
p10-123	VTN	9637.36	male	HET:PLS3het+ASO	p10	

p10-158	VTN	11166.7	male	HET:PLS3het+ASO	p10	
p10-223	VTN	14341.56	male	HET:PLS3het+ASO	p10	
p10-148	VTN	12597.95	female	HET:PLS3hom	p10	
p10-181	VTN	13840.95	female	HET:PLS3hom	p10	
p10-194	VTN	16886.92	female	HET:PLS3hom	p10	*
p10-150	VTN	13025.36	male	HET:PLS3hom	p10	
p10-152	VTN	11825.26	male	HET:PLS3hom	p10	
p10-178	VTN	12152.6	male	HET:PLS3hom	p10	
p10-122	VTN	10402.01	female	HET:PLS3hom+ASO	p10	
p10-167	VTN	12568.07	female	HET:PLS3hom+ASO	p10	
p10-244	VTN	13831.07	female	HET:PLS3hom+ASO	p10	
p10-121	VTN	9536.04	male	HET:PLS3hom+ASO	p10	
p10-126	VTN	9355.83	male	HET:PLS3hom+ASO	p10	
p10-166	VTN	11255.69	male	HET:PLS3hom+ASO	p10	
p10-201	VTN	8538.48	female	HET+ASO	p10	
p10-207	VTN	7917.79	female	HET+ASO	p10	
p10-208	VTN	8573.79	female	HET+ASO	p10	
p10-134	VTN	7333.73	male	HET+ASO	p10	
p10-135	VTN	11631.9	male	HET+ASO	p10	
p10-136	VTN	12988.8	male	HET+ASO	p10	
p10-164	VTN	10424.89	female	SMA	p10	
p10-182	VTN	7195.61	female	SMA	p10	
p10-193/3	VTN	14421.79	female	SMA	p10	*
p10-131	VTN	5784.27	male	SMA	p10	
p10-132	VTN	7391.64	male	SMA	p10	
p10-133	VTN	7433.02	male	SMA	p10	
p10-144	VTN	8348.63	female	SMA:PLS3het	p10	
p10-146	VTN	8337.9	female	SMA:PLS3het	p10	
p10-151	VTN	8217.1	female	SMA:PLS3het	p10	
p10-143	VTN	8252.64	male	SMA:PLS3het	p10	
p10-145	VTN	8497.34	male	SMA:PLS3het	p10	
p10-161	VTN	8879.06	male	SMA:PLS3het	p10	*
p10-127	VTN	8431.57	female	SMA:PLS3het+ASO	p10	
p10-155	VTN	8920.12	female	SMA:PLS3het+ASO	p10	
p10-220	VTN	8199.59	female	SMA:PLS3het+ASO	p10	
p10-154	VTN	8293.7	male	SMA:PLS3het+ASO	p10	
p10-156	VTN	7738.28	male	SMA:PLS3het+ASO	p10	
p10-219	VTN	7807.5	male	SMA:PLS3het+ASO	p10	
p10-147	VTN	5437.94	female	SMA:PLS3hom	p10	
p10-153	VTN	5053.22	female	SMA:PLS3hom	p10	
p10-196	VTN	4948.77	female	SMA:PLS3hom	p10	
p10-176	VTN	6167.02	male	SMA:PLS3hom	p10	
p10-200	VTN	5206.1	male	SMA:PLS3hom	p10	
p10-226	VTN	8133.26	male	SMA:PLS3hom	p10	*
p10-119	VTN	4295.63	female	SMA:PLS3hom+ASO	p10	
p10-218	VTN	7371.25	female	SMA:PLS3hom+ASO	p10	

p10-243	VTN	9249.35	female	SMA:PLS3hom+ASO	p10	
p10-118	VTN	4486.83	male	SMA:PLS3hom+ASO	p10	
p10-124	VTN	6419.94	male	SMA:PLS3hom+ASO	p10	
p10-125	VTN	6211.1	male	SMA:PLS3hom+ASO	p10	
p10-138	VTN	8287.45	female	SMA+ASO	p10	
p10-140	VTN	9444.19	female	SMA+ASO	p10	
p10-209	VTN	8514.9	female	SMA+ASO	p10	
p10-137	VTN	9748.56	male	SMA+ASO	p10	
p10-139	VTN	9258.67	male	SMA+ASO	p10	
p10-210	VTN	8661.2	male	SMA+ASO	p10	
p10-202	VTN	12723.9	female	WT	p10	
p10-203	VTN	10499.21	female	WT	p10	
p10-204	VTN	6913.13	female	WT	p10	
p10-205	VTN	6164.07	male	WT	p10	
p10-206	VTN	6298.81	male	WT	p10	
p10-218	VTN	6900.35	male	WT	p10	
p10-212	VTN	11166.36	female	WT+ASO	p10	
p10-213	VTN	9988.78	female	WT+ASO	p10	
p10-214	VTN	11950.55	female	WT+ASO	p10	
p10-215	VTN	9736.17	male	WT+ASO	p10	
p10-216	VTN	12063.37	male	WT+ASO	p10	
p10-217	VTN	12030.36	male	WT+ASO	p10	
p21-150	VTN	6834.78	female	HET:PLS3het+ASO	p21	
p21-163	VTN	9840.21	female	HET:PLS3het+ASO	p21	*
p21-178	VTN	6917.95	female	HET:PLS3het+ASO	p21	
p21-151	VTN	7472.47	male	HET:PLS3het+ASO	p21	
p21-157	VTN	7889.25	male	HET:PLS3het+ASO	p21	
p21-179	VTN	8116.62	male	HET:PLS3het+ASO	p21	
p21-167	VTN	7558.34	female	HET:PLS3hom+ASO	p21	
p21-211	VTN	7546.21	female	HET:PLS3hom+ASO	p21	
p21-223	VTN	6480.76	female	HET:PLS3hom+ASO	p21	
p21-158	VTN	9017.47	male	HET:PLS3hom+ASO	p21	
p21-166	VTN	7742.18	male	HET:PLS3hom+ASO	p21	
p21-190	VTN	5581.26	male	HET:PLS3hom+ASO	p21	
p21-152	VTN	7086.2	female	HET+ASO	p21	
p21-153	VTN	5733.29	female	HET+ASO	p21	
p21-175	VTN	7580.16	female	HET+ASO	p21	
p21-149	VTN	10144.71	male	HET+ASO	p21	
p21-154	VTN	6843.4	male	HET+ASO	p21	
p21-159	VTN	9910.61	male	HET+ASO	p21	
p21-156	VTN	6594.35	female	SMA:PLS3het+ASO	p21	
p21-164	VTN	6987.6	female	SMA:PLS3het+ASO	p21	
p21-168	VTN	7710.94	female	SMA:PLS3het+ASO	p21	
p21-155	VTN	7735.52	male	SMA:PLS3het+ASO	p21	
p21-165	VTN	5560.54	male	SMA:PLS3het+ASO	p21	
p21-169	VTN	7216.87	male	SMA:PLS3het+ASO	p21	

p21-196	VTN	10416.07	female	SMA:PLS3hom+ASO	p21	*
p21-221	VTN	8513.14	female	SMA:PLS3hom+ASO	p21	
p21-222	VTN	8442.92	female	SMA:PLS3hom+ASO	p21	
p21-160	VTN	6836.45	male	SMA:PLS3hom+ASO	p21	*
p21-162	VTN	8478.1	male	SMA:PLS3hom+ASO	p21	
p21-212	VTN	8106.19	male	SMA:PLS3hom+ASO	p21	
p21-182	VTN	7729.77	female	SMA+ASO	p21	
p21-183	VTN	6719.95	female	SMA+ASO	p21	
p21-189	VTN	9007.67	female	SMA+ASO	p21	
p21-184	VTN	6998.05	male	SMA+ASO	p21	
p21-185	VTN	6358.44	male	SMA+ASO	p21	
p21-191	VTN	5491.48	male	SMA+ASO	p21	
p21-207	VTN	6935.94	female	WT+ASO	p21	
p21-214	VTN	7637.69	female	WT+ASO	p21	
p21-215	VTN	6575.61	female	WT+ASO	p21	
p21-197	VTN	7210.6	male	WT+ASO	p21	
p21-198	VTN	8696.04	male	WT+ASO	p21	
p21-208	VTN	7754.48	male	WT+ASO	p21	

8. Publications

8.1. Scientific publications

2022: Scientific publication as author
Strathmann EA, Hölker I, Tschernoster N, Hosseinibarkooie S, Come J, Martinat C, Altmüller J, Brunhilde Wirth B
Epigenetic regulation of plastin 3 expression by the macrosatellite DXZ4 and the transcriptional regulator CHD4
Under revision at AJHG

2021: Review as shared author
Wolff L, **Strathmann EA**, Müller I, Mählich D, Veltmann C, Niehoff A, Wirth B
Plastin 3 in health and disease: a matter of balance
Cell Mol Life Sci. 2021 Jul;78(13):5275-5301.doi: 10.1007/s00018-021-03843-5. Epub 2021 May 23.

2018: Scientific publication as author
Strathmann EA, Peters M, Hosseinibarkooie S, Rigo FW, Bennett CF, Zaworski PG, Chen KS, Nothnagel M, Wirth B
Evaluation of potential effects of Plastin 3 overexpression and low-dose SMN-antisense oligonucleotides on putative biomarkers in spinal muscular atrophy mice.
PLoS One. 2018 Sep 6;13(9):e0203398. doi: 10.1371/journal.pone.0203398. eCollection 2018.

2018: Scientific publication as co-author
Karakaya M, Storbeck M, **Strathmann EA**, Delle Vedove A, Hölker I, Altmueller J, Naghiyeva L, Schmitz-Steinkrüger L, Vezyroglou K, Motameny S, Alawbathani S, Thiele H, Polat AI, Okur D, Boostani R, Karimiani EG[5], Wunderlich G, Ardicli D, Topaloglu H, Kirschner J, Schrank B, Maroofian R, Magnusson O, Yis U, Nürnberg P, Heller R, Wirth B
Targeted sequencing with expanded gene profile enables high diagnostic yield in non-5q-spinal muscular atrophies.
Hum Mutat. 2018 Sep;39(9):1284-1298. doi: 10.1002/humu.23560. Epub 2018 Jul 25.

2015: Scientific publication as co-author
Petersen C, Saebelfeld M, Barbosa C, Pees B, Hermann RJ, Schalkowski R, **Strathmann EA**, Dirksen P, Schulenburg H
Ten years of life in compost: temporal and spatial variation of North German *Caenorhabditis elegans* populations.
Ecol Evol. 2015 Aug;5(16):3250-63. doi: 10.1002/ece3.1605. Epub 2015 Jul 15.

2014: Scientific publication as co-author
Petersen C, Dirksen P, Prahl S, Strathmann EA, Schulenburg H
The prevalence of *Caenorhabditis elegans* across 1.5 years in selected North German locations: the importance of substrate type, abiotic parameters, and Caenorhabditis competitors.
BMC Ecol. 2014 Feb 6;14:4.doi: 10.1186/1472-6785-14-4.

9. References

1. Ackermann, B., Krober, S., Torres-Benito, L., Borgmann, A., Peters, M., Hosseini Barkooie, S.M., Tejero, R., Jakubik, M., Schreml, J., Milbradt, J.*, et al.* (2013). Plastin 3 ameliorates spinal muscular atrophy via delayed axon pruning and improves neuromuscular junction functionality. Hum Mol Genet *22*, 1328-1347.

2. Agarwal, P., Zwolanek, D., Keene, D.R., Schulz, J.N., Blumbach, K., Heinegard, D., Zaucke, F., Paulsson, M., Krieg, T., Koch, M.*, et al.* (2012). Collagen XII and XIV, new partners of cartilage oligomeric matrix protein in the skin extracellular matrix suprastructure. J Biol Chem *287*, 22549-22559.

3. Alrafiah, A., Karyka, E., Coldicott, I., Iremonger, K., Lewis, K.E., Ning, K., and Azzouz, M. (2018). Plastin 3 Promotes Motor Neuron Axonal Growth and Extends Survival in a Mouse Model of Spinal Muscular Atrophy. Mol Ther Methods Clin Dev *9*, 81-89.

4. Alves, C.R.R., Zhang, R., Johnstone, A.J., Garner, R., Nwe, P.H., Siranosian, J.J., and Swoboda, K.J. (2020). Serum creatinine is a biomarker of progressive denervation in spinal muscular atrophy. Neurology *94*, e921-e931.

5. Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data (Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom).

6. Arellano-Garcia, M.E., Li, R., Liu, X., Xie, Y., Yan, X., Loo, J.A., and Hu, S. (2010). Identification of tetranectin as a potential biomarker for metastatic oral cancer. Int J Mol Sci *11*, 3106-3121.

7. Arends, T., Dege, C., Bortnick, A., Danhorn, T., Knapp, J.R., Jia, H., Harmacek, L., Fleenor, C.J., Straign, D., Walton, K.*, et al.* (2019). CHD4 is essential for transcriptional repression and lineage progression in B lymphopoiesis. Proc Natl Acad Sci U S A *116*, 10927-10936.

8. Arnold, W.D., Duque, S., Iyer, C.C., Zaworski, P., McGovern, V.L., Taylor, S.J., von Herrmann, K.M., Kobayashi, D.T., Chen, K.S., Kolb, S.J.*, et al.* (2016). Normalization of Patient-Identified Plasma Biomarkers in SMNDelta7 Mice following Postnatal SMN Restoration. PLoS One *11*, e0167077.

9. Arnold, W.D., Porensky, P.N., McGovern, V.L., Iyer, C.C., Duque, S., Li, X., Meyer, K., Schmelzer, L., Kaspar, B.K., Kolb, S.J.*, et al.* (2014). Electrophysiological Biomarkers in Spinal Muscular Atrophy: Preclinical Proof of Concept. Ann Clin Transl Neurol *1*, 34-44.

10. Azukizawa, M., Ito, H., Hamamoto, Y., Fujii, T., Morita, Y., Okahata, A., Tomizawa, T., Furu, M., Nishitani, K., Kuriyama, S.*, et al.* (2018). The Effects of Well-Rounded Exercise Program on Systemic Biomarkers Related to Cartilage Metabolism. Cartilage, 1947603518767998.

11. Balaton, B.P., Cotton, A.M., and Brown, C.J. (2015). Derivation of consensus inactivation status for X-linked genes from genome-wide studies. Biol Sex Differ *6*, 35.

12. Bedford, M.T., and van Helden, P.D. (1987). Hypomethylation of DNA in pathological conditions of the human prostate. Cancer Res *47*, 5274-5276.

13. Berletch, J.B., Ma, W., Yang, F., Shendure, J., Noble, W.S., Disteche, C.M., and Deng, X. (2015). Escape from X inactivation varies in mouse tissues. PLoS Genet *11*, e1005079.

14. Bernabo, P., Tebaldi, T., Groen, E.J.N., Lane, F.M., Perenthaler, E., Mattedi, F., Newbery, H.J., Zhou, H., Zuccotti, P., Potrich, V.*, et al.* (2017). In Vivo Translatome Profiling in Spinal Muscular Atrophy Reveals a Role for SMN Protein in Ribosome Biology. Cell Rep *21*, 953-965.

15. Bertelsen, B., Tumer, Z., and Ravn, K. (2011). Three new loci for determining x chromosome inactivation patterns. J Mol Diagn *13*, 537-540.

16. Bonanno, S., Marcuzzo, S., Malacarne, C., Giagnorio, E., Masson, R., Zanin, R., Arnoldi, M.T., Andreetta, F., Simoncini, O., Venerando, A.*, et al.* (2020). Circulating MyomiRs as Potential Biomarkers to Monitor Response to Nusinersen in Pediatric SMA Patients. Biomedicines *8*.

17. Bonati, U., Holiga, S., Hellbach, N., Risterucci, C., Bergauer, T., Tang, W., Hafner, P., Thoeni, A., Bieri, O., Gerlach, I.*, et al.* (2017). Longitudinal characterization of biomarkers for spinal muscular atrophy. Ann Clin Transl Neurol *4*, 292-304.

18. Bonora, G., Deng, X., Fang, H., Ramani, V., Qiu, R., Berletch, J.B., Filippova, G.N., Duan, Z., Shendure, J., Noble, W.S.*, et al.* (2018). Orientation-dependent Dxz4 contacts shape the 3D structure of the inactive X chromosome. Nature communications *9*, 1445.

19. Bornelov, S., Reynolds, N., Xenophontos, M., Gharbi, S., Johnstone, E., Floyd, R., Ralser, M., Signolet, J., Loos, R., Dietmann, S.*, et al.* (2018). The Nucleosome Remodeling and Deacetylation Complex Modulates Chromatin Structure at Sites of Active Transcription to Fine-Tune Gene Expression. Mol Cell *71*, 56-72 e54.

20. Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol *34*, 525-527.

21. Bretscher, A. (1981). Fimbrin is a cytoskeletal protein that crosslinks F-actin in vitro. Proc Natl Acad Sci U S A *78*, 6849-6853.

22. Bretscher, A., and Weber, K. (1980). Fimbrin, a new microfilament-associated protein present in microvilli and other cell surface structures. J Cell Biol *86*, 335-340.

23. Briggs, M.D., Hoffman, S.M., King, L.M., Olsen, A.S., Mohrenweiser, H., Leroy, J.G., Mortier, G.R., Rimoin, D.L., Lachman, R.S., Gaines, E.S.*, et al.* (1995). Pseudoachondroplasia and multiple epiphyseal dysplasia due to mutations in the cartilage oligomeric matrix protein gene. Nat Genet *10*, 330-336.

24. Brothman, A.R., Swanson, G., Maxwell, T.M., Cui, J., Murphy, K.J., Herrick, J., Speights, V.O., Isaac, J., and Rohr, L.R. (2005). Global hypomethylation is common in prostate cancer cells: a quantitative predictor for clinical outcome? Cancer genetics and cytogenetics *156*, 31-36.

25. Brylka, L.J., Koppert, S., Babler, A., Kratz, B., Denecke, B., Yorgan, T.A., Etich, J., Costa, I.G., Brachvogel, B., Boor, P.*, et al.* (2017). Post-weaning epiphysiolysis causes distal femur dysplasia and foreshortened hindlimbs in fetuin-A-deficient mice. PLoS One *12*, e0187030.

26. Brzustowicz, L.M., Lehner, T., Castilla, L.H., Penchaszadeh, G.K., Wilhelmsen, K.C., Daniels, R., Davies, K.E., Leppert, M., Ziter, F., Wood, D.*, et al.* (1990). Genetic mapping of chronic childhood-onset spinal muscular atrophy to chromosome 5q11.2-13.3. Nature *344*, 540-541.

27. Burglen, L., Lefebvre, S., Clermont, O., Burlet, P., Viollet, L., Cruaud, C., Munnich, A., and Melki, J. (1996). Structure and organization of the human survival motor neurone (SMN) gene. Genomics *32*, 479-482.

28. Burnett, B.G., Munoz, E., Tandon, A., Kwon, D.Y., Sumner, C.J., and Fischbeck, K.H. (2009). Regulation of SMN protein stability. Mol Cell Biol *29*, 1107-1115.

29. Butchbach, M.E.R. (2021). Genomic Variability in the Survival Motor Neuron Genes (SMN1 and SMN2): Implications for Spinal Muscular Atrophy Phenotype and Therapeutics Development. Int J Mol Sci *22*.

30. Cai, Y., Geutjes, E.J., de Lint, K., Roepman, P., Bruurs, L., Yu, L.R., Wang, W., van Blijswijk, J., Mohammad, H., de Rink, I.*, et al.* (2014). The NuRD complex cooperates with DNMTs to maintain silencing of key colorectal tumor suppressor genes. Oncogene *33*, 2157-2168.

31. Carrel, L., and Willard, H.F. (2005). X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature *434*, 400-404.

32. Cartegni, L., and Krainer, A.R. (2002). Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. Nat Genet *30*, 377-384.

33. Catapano, F., Zaharieva, I., Scoto, M., Marrosu, E., Morgan, J., Muntoni, F., and Zhou, H. (2016). Altered Levels of MicroRNA-9, -206, and -132 in Spinal Muscular Atrophy and Their Response to Antisense Oligonucleotide Therapy. Molecular therapy Nucleic acids *5*, e331.

34. Chadwick, B.P. (2008). DXZ4 chromatin adopts an opposing conformation to that of the surrounding chromosome and acquires a novel inactive X-specific role involving CTCF and antisense transcripts. Genome Res *18*, 1259-1269.

35. Chafel, M.M., Shen, W., and Matsudaira, P. (1995). Sequential expression and differential localization of I-, L-, and T-fimbrin during differentiation of the mouse intestine and yolk sac. Dev Dyn *203*, 141-151.

36. Chang, C.L., Huang, C.R., Chang, S.J., Wu, C.C., Chen, H.H., Luo, C.W., and Yip, H.K. (2021). CHD4 as an important mediator in regulating the malignant behaviors of colorectal cancer. Int J Biol Sci *17*, 1660-1670.

37. Chaytow, H., Huang, Y.T., Gillingwater, T.H., and Faller, K.M.E. (2018). The role of survival motor neuron protein (SMN) in protein homeostasis. Cell Mol Life Sci *75*, 3877-3894.

38. Chen, H., Li, H., Zhao, J., Peng, P., Shao, M., Wu, H., Wang, X., Chen, L., Zhang, Q., Ruan, Y.*, et al.* (2017). High Intratumoral Expression of Tetranectin Associates with Poor Prognosis of Patients with Gastric Cancer after Gastrectomy. J Cancer *8*, 3623-3630.

39. Chen, H.Y., Chiu, Y.L., Hsu, S.P., Pai, M.F., Yang, J.Y., and Peng, Y.S. (2013). Low serum fetuin A levels and incident stroke in patients with maintenance haemodialysis. Eur J Clin Invest *43*, 387-396.

40. Chen, T.H. (2020). Circulating microRNAs as potential biomarkers and therapeutic targets in spinal muscular atrophy. Therapeutic advances in neurological disorders *13*, 1756286420979954.

41. Chudnovsky, Y., Kim, D., Zheng, S., Whyte, W.A., Bansal, M., Bray, M.A., Gopal, S., Theisen, M.A., Bilodeau, S., Thiru, P.*, et al.* (2014). ZFHX4 interacts with the NuRD core member CHD4 and regulates the glioblastoma tumor-initiating cell state. Cell Rep *6*, 313-324.

42. Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szczesniak, M.W., Gaffney, D.J., Elo, L.L., Zhang, X.*, et al.* (2016). A survey of best practices for RNA-seq data analysis. Genome Biol *17*, 13.

43. Coovert, D.D., Le, T.T., McAndrew, P.E., Strasswimmer, J., Crawford, T.O., Mendell, J.R., Coulson, S.E., Androphy, E.J., Prior, T.W., and Burghes, A.H. (1997). The survival motor neuron protein in spinal muscular atrophy. Hum Mol Genet *6*, 1205-1214.

44. Cotton, A.M., Ge, B., Light, N., Adoue, V., Pastinen, T., and Brown, C.J. (2013). Analysis of expressed SNPs identifies variable extents of expression from the human inactive X chromosome. Genome Biol *14*, R122.

45. Cotton, A.M., Price, E.M., Jones, M.J., Balaton, B.P., Kobor, M.S., and Brown, C.J. (2015). Landscape of DNA methylation on the X chromosome reflects CpG density, functional chromatin state and X-chromosome inactivation. Hum Mol Genet *24*, 1528-1539.

46. Crawford, T.O., Paushkin, S.V., Kobayashi, D.T., Forrest, S.J., Joyce, C.L., Finkel, R.S., Kaufmann, P., Swoboda, K.J., Tiziano, D., Lomastro, R.*, et al.* (2012). Evaluation of SMN protein, transcript, and copy number in the biomarkers for spinal muscular atrophy (BforSMA) clinical study. PLoS One *7*, e33572.

47. Czech, C., Tang, W., Bugawan, T., Mano, C., Horn, C., Iglesias, V.A., Frohner, S., Zaworski, P.G., Paushkin, S., Chen, K.*, et al.* (2015). Biomarker for Spinal Muscular Atrophy: Expression of SMN in Peripheral Blood of SMA Patients and Healthy Controls. PLoS One *10*, e0139950.

48. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M.*, et al.* (2021). Twelve years of SAMtools and BCFtools. Gigascience *10*.

49. Darras, B.T., Crawford, T.O., Finkel, R.S., Mercuri, E., De Vivo, D.C., Oskoui, M., Tizzano, E.F., Ryan, M.M., Muntoni, F., Zhao, G.*, et al.* (2019). Neurofilament as a potential biomarker for spinal muscular atrophy. Ann Clin Transl Neurol *6*, 932-944.

50. Darrow, E.M., Huntley, M.H., Dudchenko, O., Stamenova, E.K., Durand, N.C., Sun, Z., Huang, S.C., Sanborn, A.L., Machol, I., Shamim, M.*, et al.* (2016). Deletion of DXZ4 on the human inactive X chromosome alters higher-order genome architecture. Proc Natl Acad Sci U S A *113*, E4504-4512.

51. de Capoa, A., Musolino, A., Della Rosa, S., Caiafa, P., Mariani, L., Del Nonno, F., Vocaturo, A., Donnorso, R.P., Niveleau, A., and Grappelli, C. (2003). DNA demethylation is directly related to tumour progression: evidence in normal, pre-malignant and malignant cells from uterine cervix samples. Oncol Rep *10*, 545-549.

52. Denslow, S.A., and Wade, P.A. (2007). The human Mi-2/NuRD complex and gene regulation. Oncogene *26*, 5433-5438.

53. Dimitriadi, M., Derdowski, A., Kalloo, G., Maginnis, M.S., O'Hern, P., Bliska, B., Sorkac, A., Nguyen, K.C., Cook, S.J., Poulogiannis, G.*, et al.* (2016). Decreased function of survival motor neuron protein impairs endocytic pathways. Proc Natl Acad Sci U S A *113*, E4377-4386.

54. Dunn, O.J. (1958). Estimation of the Means of Dependent-Variables. Ann Math Stat *29*, 1095-1111.

55. Dunn, O.J. (1961). Multiple Comparisons among Means. J Am Stat Assoc *56*, 52-&.

56. Dunnett, C.W. (1955). A multiple comparison procedure for comparing several treatments with a control. J Am Stat Assoc *50*, 1096-1121.

57. Ehrlich, M. (2002). DNA methylation in cancer: too much, but also too little. Oncogene *21*, 5400-5413.

58. Eichelberger, E.J., Alves, C.R.R., Zhang, R., Petrillo, M., Cullen, P., Farwell, W., Hurt, J.A., Staropoli, J.F., and Swoboda, K.J. (2021). Increased systemic HSP70B levels in spinal muscular atrophy infants. Ann Clin Transl Neurol *8*, 1495-1501.

59. El Defrawy, A.O., Gheita, T.A., Raslan, H.M., El Ansary, M.M., and El Awar, A.H. (2016). Serum and synovial cartilage oligomeric matrix protein levels in early and established rheumatoid arthritis. Z Rheumatol *75*, 917-923.

60. Fabregat, A., Sidiropoulos, K., Garapati, P., Gillespie, M., Hausmann, K., Haw, R., Jassal, B., Jupe, S., Korninger, F., McKay, S.*, et al.* (2016). The Reactome pathway Knowledgebase. Nucleic Acids Res *44*, D481-487.

61. Fedou, C., Camus, M., Lescat, O., Feuillet, G., Mueller, I., Ross, B., Buleon, M., Neau, E., Alves, M., Goudouneche, D.*, et al.* (2021). Mapping of the amniotic fluid proteome of fetuses with congenital anomalies of the kidney and urinary tract identifies plastin 3 as a protein involved in glomerular integrity. J Pathol *254*, 575-588.

62. Feldkotter, M., Schwarzer, V., Wirth, R., Wienker, T.F., and Wirth, B. (2002). Quantitative analyses of SMN1 and SMN2 based on real-time lightCycler PCR: fast and highly reliable carrier testing and prediction of severity of spinal muscular atrophy. Am J Hum Genet *70*, 358-368.

63. Figueroa, D.M., Darrow, E.M., and Chadwick, B.P. (2015). Two novel DXZ4-associated long noncoding RNAs show developmental changes in expression coincident with heterochromatin formation at the human (Homo sapiens) macrosatellite repeat. Chromosome Res *23*, 733-752.

64. Finkel, R.S., Crawford, T.O., Swoboda, K.J., Kaufmann, P., Juhasz, P., Li, X., Guo, Y., Li, R.H., Trachtenberg, F., Forrest, S.J.*, et al.* (2012). Candidate proteins, metabolites and transcripts in the Biomarkers for Spinal Muscular Atrophy (BforSMA) clinical study. PLoS One *7*, e35462.

65. Gama-Sosa, M.A., Slagel, V.A., Trewyn, R.W., Oxenhandler, R., Kuo, K.C., Gehrke, C.W., and Ehrlich, M. (1983). The 5-methylcytosine content of DNA from human tumors. Nucleic Acids Res *11*, 6883-6894.

66. Geary, R.S., Leeds, J.M., Henry, S.P., Monteith, D.K., and Levin, A.A. (1997). Antisense oligonucleotide inhibitors for the treatment of cancer: 1. Pharmacokinetic properties of phosphorothioate oligodeoxynucleotides. Anticancer Drug Des *12*, 383-393.

67. Geary, R.S., Yu, R.Z., Watanabe, T., Henry, S.P., Hardee, G.E., Chappell, A., Matson, J., Sasmor, H., Cummins, L., and Levin, A.A. (2003). Pharmacokinetics of a tumor necrosis factor-alpha phosphorothioate 2'-O-(2-methoxyethyl) modified antisense oligonucleotide: comparison across species. Drug Metab Dispos *31*, 1419-1428.

68. Geens, M., and Chuva De Sousa Lopes, S.M. (2017). X chromosome inactivation in human pluripotent stem cells as a model for human development: back to the drawing board? Hum Reprod Update *23*, 520-532.

69. Gerdhem, P., Topalis, C., Grauers, A., Stubendorff, J., Ohlin, A., and Karlsson, K.M. (2015). Serum level of cartilage oligomeric matrix protein is lower in children with idiopathic scoliosis than in non-scoliotic controls. Eur Spine J *24*, 256-261.

70. Giacalone, J., Friedes, J., and Francke, U. (1992). A novel GC-rich human macrosatellite VNTR in Xq24 is differentially methylated on active and inactive X chromosomes. Nat Genet *1*, 137-143.

71. Giavazzi, A., Setola, V., Simonati, A., and Battaglia, G. (2006). Neuronal-specific roles of the survival motor neuron protein: evidence from survival motor neuron expression patterns in the developing human central nervous system. J Neuropathol Exp Neurol *65*, 267-277.

72. Girden, E.R. (1992). ANOVA: Repeated Measures (SAGE Publications).

73. Goldstein, M.E., Cooper, H.S., Bruce, J., Carden, M.J., Lee, V.M., and Schlaepfer, W.W. (1987). Phosphorylation of neurofilament proteins and chromatolysis following transection of rat sciatic nerve. J Neurosci *7*, 1586-1594.

74. Hao, C., Cui, Y., Owen, S., Li, W., Cheng, S., and Jiang, W.G. (2017). Human osteopontin: Potential clinical applications in cancer (Review). Int J Mol Med *39*, 1327-1337.

75. Haramati, S., Chapnik, E., Sztainberg, Y., Eilam, R., Zwang, R., Gershoni, N., McGlinn, E., Heiser, P.W., Wills, A.M., Wirguin, I.*, et al.* (2010). miRNA malfunction causes spinal motor neuron disease. Proc Natl Acad Sci U S A *107*, 13111-13116.

76. Hawley, Z.C.E., Campos-Melo, D., Droppelmann, C.A., and Strong, M.J. (2017). MotomiRs: miRNAs in Motor Neuron Function and Disease. Front Mol Neurosci *10*, 127.

77. Haylock, D.N., and Nilsson, S.K. (2006). Osteopontin: a bridge between bone and blood. Br J Haematol *134*, 467-474.

78. Hecht, J.T., Nelson, L.D., Crowder, E., Wang, Y., Elder, F.F., Harrison, W.R., Francomano, C.A., Prange, C.K., Lennon, G.G., Deere, M.*, et al.* (1995). Mutations in exon 17B of cartilage oligomeric matrix protein (COMP) cause pseudoachondroplasia. Nat Genet *10*, 325-329.

79. Heesen, L., Peitz, M., Dimos, J.T., Brustle, O., and Wirth, B. (2010). Generation and application of iPS cells & motoneurons for tissue-specific analysis of SMN-specific changes in SMA patients. Transgenic Res *19*, 332.

80. Heesen, L., Peitz, M., Torres-Benito, L., Holker, I., Hupperich, K., Dobrindt, K., Jungverdorben, J., Ritzenhofen, S., Weykopf, B., Eckert, D.*, et al.* (2016). Plastin 3 is upregulated in iPSC-derived motoneurons from asymptomatic SMN1-deleted individuals. Cell Mol Life Sci *73*, 2089-2104.

81. Helmken, C., Hofmann, Y., Schoenen, F., Oprea, G., Raschke, H., Rudnik-Schoneborn, S., Zerres, K., and Wirth, B. (2003). Evidence for a modifying pathway in SMA discordant families: reduced SMN level decreases the amount of its interacting partners and Htra2-beta1. Hum Genet *114*, 11-21.

82. Hendrich, B., and Bird, A. (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins. Mol Cell Biol *18*, 6538-6547.

83. Hoffmeister, H., Fuchs, A., Erdel, F., Pinz, S., Grobner-Ferreira, R., Bruckmann, A., Deutzmann, R., Schwartz, U., Maldonado, R., Huber, C.*, et al.* (2017). CHD3 and CHD4 form distinct NuRD complexes with different yet overlapping functionality. Nucleic Acids Res *45*, 10534-10554.

84. Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. Scand J Stat *6*, 65-70.

85. Hosseinibarkooie, S., Peters, M., Torres-Benito, L., Rastetter, R.H., Hupperich, K., Hoffmann, A., Mendoza-Ferreira, N., Kaczmarek, A., Janzen, E., Milbradt, J.*, et al.* (2016). The Power of Human Protective Modifiers: PLS3 and CORO1C Unravel Impaired Endocytosis in Spinal Muscular Atrophy and Rescue SMA Phenotype. Am J Hum Genet *99*, 647-665.

86. Hsieh-Li, H.M., Chang, J.G., Jong, Y.J., Wu, M.H., Wang, N.M., Tsai, C.H., and Li, H. (2000). A mouse model for spinal muscular atrophy. Nat Genet *24*, 66-70.

87. Hua, Y., Sahashi, K., Rigo, F., Hung, G., Horev, G., Bennett, C.F., and Krainer, A.R. (2011). Peripheral SMN restoration is essential for long-term rescue of a severe spinal muscular atrophy mouse model. Nature *478*, 123-126.

88. Iyer, C.C., Wang, X., Renusch, S.R., Duque, S.I., Wehr, A.M., Mo, X.M., McGovern, V.L., Arnold, W.D., Burghes, A.H., and Kolb, S.J. (2017). SMN Blood Levels in a Porcine Model of Spinal Muscular Atrophy. J Neuromuscul Dis *4*, 59-66.

89. Janzen, E., Wolff, L., Mendoza-Ferreira, N., Hupperich, K., Delle Vedove, A., Hosseinibarkooie, S., Kye, M.J., and Wirth, B. (2019). PLS3 Overexpression Delays Ataxia in Chp1 Mutant Mice. Frontiers in neuroscience *13*, 993.

90. Jiao, Q., Wei, L., Chen, C., Li, P., Wang, X., Li, Y., Guo, L., Zhang, C., and Wei, X. (2016). Cartilage oligomeric matrix protein and hyaluronic acid are sensitive serum biomarkers for early cartilage lesions in the knee joint. Biomarkers *21*, 146-151.

91. Jones, C.L., Ferreira, S., McKenzie, R.C., Tosi, I., Caesar, J.A., Bagot, M., Whittaker, S.J., and Mitchell, T.J. (2012). Regulation of T-plastin expression by promoter hypomethylation in primary cutaneous T-cell lymphoma. J Invest Dermatol *132*, 2042-2049.

92. Kaifer, K.A., Villalon, E., Osman, E.Y., Glascock, J.J., Arnold, L.L., Cornelison, D.D.W., and Lorson, C.L. (2017). Plastin-3 extends survival and reduces severity in mouse models of spinal muscular atrophy. JCI Insight *2*, e89970.

93. Karampela, I., Kandri, E., Antonakos, G., Vogiatzakis, E., Christodoulatos, G.S., Nikolaidou, A., Dimopoulos, G., Armaganidis, A., and Dalamaga, M. (2017). Kinetics of circulating fetuin-A may predict mortality independently from adiponectin, high molecular weight adiponectin and prognostic factors in critically ill patients with sepsis: A prospective study. J Crit Care *41*, 78-85.

94. Karpova, T.S., Tatchell, K., and Cooper, J.A. (1995). Actin filaments in yeast are unstable in the absence of capping protein or fimbrin. J Cell Biol *131*, 1483-1493.

95. Khatri, I.A., Chaudhry, U.S., Seikaly, M.G., Browne, R.H., and Iannaccone, S.T. (2008). Low bone mineral density in spinal muscular atrophy. J Clin Neuromuscul Dis *10*, 11-17.

96. Kiedrowski, L.A., Raca, G., Laffin, J.J., Nisler, B.S., Leonhard, K., McIntire, E., and Mongomery, K.D. (2011). DNA methylation assay for X-chromosome inactivation in female human iPS cells. Stem Cell Rev Rep *7*, 969-975.

97. Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol *37*, 907-915.

98. Kitamura, Y., Usami, R., Ichihara, S., Kida, H., Satoh, M., Tomimoto, H., Murata, M., and Oikawa, S. (2017). Plasma protein profiling for potential biomarkers in the early diagnosis of Alzheimer's disease. Neurol Res *39*, 231-238.

99. Kluzek, S., Bay-Jensen, A.C., Judge, A., Karsdal, M.A., Shorthose, M., Spector, T., Hart, D., Newton, J.L., and Arden, N.K. (2015). Serum cartilage oligomeric matrix protein and

development of radiographic and painful knee osteoarthritis. A community-based cohort of middle-aged women. Biomarkers *20*, 557-564.

100. Kobayashi, D.T., Shi, J., Stephen, L., Ballard, K.L., Dewey, R., Mapes, J., Chung, B., McCarthy, K., Swoboda, K.J., Crawford, T.O.*, et al.* (2013). SMA-MAP: a plasma protein panel for spinal muscular atrophy. PLoS One *8*, e60113.

101. Kolb, S.J., Coffey, C.S., Yankey, J.W., Krosschell, K., Arnold, W.D., Rutkove, S.B., Swoboda, K.J., Reyna, S.P., Sakonju, A., Darras, B.T.*, et al.* (2016). Baseline results of the NeuroNEXT spinal muscular atrophy infant biomarker study. Ann Clin Transl Neurol *3*, 132-145.

102. Kong, L., Wang, X., Choe, D.W., Polley, M., Burnett, B.G., Bosch-Marce, M., Griffin, J.W., Rich, M.M., and Sumner, C.J. (2009). Impaired synaptic vesicle release and immaturity of neuromuscular junctions in spinal muscular atrophy mice. J Neurosci *29*, 842-851.

103. Kujawski, R., Przybylowska-Sygut, K., Mik, M., Lewandowski, M., Trzcinski, R., Berut, M., Dziki, L., Majsterek, I., and Dziki, A. (2015). Expression of the PLS3 Gene in Circulating Cells in Patients with Colorectal Cancer. Pol Przegl Chir *87*, 59-64.

104. Kuraoka, M., Kimura, E., Nagata, T., Okada, T., Aoki, Y., Tachimori, H., Yonemoto, N., Imamura, M., and Takeda, S. (2016). Serum Osteopontin as a Novel Biomarker for Muscle Regeneration in Duchenne Muscular Dystrophy. Am J Pathol *186*, 1302-1312.

105. Kurashige, J., Yokobori, T., Mima, K., Sawada, G., Takahashi, Y., Ueo, H., Takano, Y., Matsumura, T., Uchi, R., Eguchi, H.*, et al.* (2019). Plastin3 is associated with epithelial-mesenchymal transition and poor prognosis in gastric cancer. Oncol Lett *17*, 2393-2399.

106. Kuriyama, K., Yokobori, T., Sohda, M., Nakazawa, N., Yajima, T., Naruse, I., Kuwano, H., Shirabe, K., Kaira, K., and Saeki, H. (2021). Plasma plastin-3: A tumor marker in patients with non-small-cell lung cancer treated with nivolumab. Oncol Lett *21*, 11.

107. Larocca, C., and Kupper, T. (2019). Mycosis Fungoides and Sezary Syndrome: An Update. Hematol Oncol Clin North Am *33*, 103-120.

108. Lebensztejn, D.M., Bialokoz-Kalinowska, I., Klusek-Oksiuta, M., Tarasow, E., Wojtkowska, M., and Kaczmarski, M. (2014). Serum fetuin A concentration is elevated in children with non-alcoholic fatty liver disease. Adv Med Sci *59*, 81-84.

109. Lee, S., Lee, S., Ouellette, S., Park, W.Y., Lee, E.A., and Park, P.J. (2017). NGSCheckMate: software for validating sample identity in next-generation sequencing studies within and across data types. Nucleic Acids Res *45*, e103.

110. Lefebvre, S., Burglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., Benichou, B., Cruaud, C., Millasseau, P., Zeviani, M.*, et al.* (1995). Identification and characterization of a spinal muscular atrophy-determining gene. Cell *80*, 155-165.

111. Lefebvre, S., Burlet, P., Liu, Q., Bertrandy, S., Clermont, O., Munnich, A., Dreyfuss, G., and Melki, J. (1997). Correlation between severity and SMN protein level in spinal muscular atrophy. Nat Genet *16*, 265-269.

112. Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics *27*, 2987-2993.

113. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754-1760.

114. Li, H., Ruan, J., and Durbin, R. (2008). Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res *18*, 1851-1858.

115. Lin, C.S., Lau, A., Huynh, T., and Lue, T.F. (1999). Differential regulation of human T-plastin gene in leukocytes and non-leukocytes: identification of the promoter, enhancer, and CpG island. DNA Cell Biol *18*, 27-37.

116. Lin, C.S., Park, T., Chen, Z.P., and Leavitt, J. (1993). Human plastin genes. Comparative gene structure, chromosome location, and differential expression in normal and neoplastic cells. J Biol Chem *268*, 2781-2792.

117. Liu, L., Zhao, M., Xie, Z.G., Liu, J., Peng, H.P., Pei, Y.F., Sun, H.P., and Zhang, L. (2020). Twelve New Genomic Loci Associated With Bone Mineral Density. Front Endocrinol (Lausanne) *11*, 243.

118. Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402-408.

119. Lorson, C.L., Hahnen, E., Androphy, E.J., and Wirth, B. (1999). A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. Proc Natl Acad Sci U S A *96*, 6307-6311.

120. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.

121. Lyberopoulou, A., Aravantinos, G., Efstathopoulos, E.P., Nikiteas, N., Bouziotis, P., Isaakidou, A., Papalois, A., Marinos, E., and Gazouli, M. (2015). Mutational analysis of circulating tumor cells from colorectal cancer patients and correlation with primary tumor tissue. PLoS One *10*, e0123902.

122. Magen, I., Aharoni, S., Yacovzada, N.S., Tokatly Latzer, I., Alves, C.R.R., Sagi, L., Fattal-Valevski, A., Swoboda, K.J., Katz, J., Bruckheimer, E.*, et al.* (2022). Muscle microRNAs in the cerebrospinal fluid predict clinical response to nusinersen therapy in type II and type III spinal muscular atrophy patients. Eur J Neurol *29*, 2420-2430.

123. Magri, F., Vanoli, F., and Corti, S. (2018). miRNA in spinal muscular atrophy pathogenesis and therapy. Journal of cellular and molecular medicine *22*, 755-767.

124. Mählich, D., Glasmacher, A., Müller, I., Oppermann, J., Grevenstein, D., Eysel, P., Heilig, J., Wirth, B., Zaucke, F., and Niehoff, A. (2021). Expression and localization of thrombospondins, plastin 3, and STIM1 in different cartilage compartments of the osteoarthritic varus knee. Int J Mol Sci *22*, 1-19.

125. Main, M., Kairon, H., Mercuri, E., and Muntoni, F. (2003). The Hammersmith functional motor scale for children with spinal muscular atrophy: a scale to test ability and monitor progress in children with limited ambulation. Eur J Paediatr Neurol *7*, 155-159.

126. Makitie, R.E., Hackl, M., Weigl, M., Frischer, A., Kampe, A., Costantini, A., Grillari, J., and Makitie, O. (2020a). Unique, Gender-Dependent Serum microRNA Profile in PLS3 Gene-Related Osteoporosis. J Bone Miner Res.

127. Makitie, R.E., Kampe, A., Costantini, A., Alm, J.J., Magnusson, P., and Makitie, O. (2020b). Biomarkers in WNT1 and PLS3 Osteoporosis: Altered Concentrations of DKK1 and FGF23. J Bone Miner Res *35*, 901-912.

128. Makitie, R.E., Niinimaki, T., Suo-Palosaari, M., Kampe, A., Costantini, A., Toiviainen-Salo, S., Niinimaki, J., and Makitie, O. (2020c). PLS3 Mutations Cause Severe Age and Sex-Related Spinal Pathology. Front Endocrinol (Lausanne) *11*, 393.

129. Mansson, B., Carey, D., Alini, M., Ionescu, M., Rosenberg, L.C., Poole, A.R., Heinegard, D., and Saxne, T. (1995). Cartilage and bone metabolism in rheumatoid arthritis. Differences between rapid and slow progression of disease identified by serum markers of cartilage metabolism. J Clin Invest *95*, 1071-1077.

130. Marfella, C.G., and Imbalzano, A.N. (2007). The Chd family of chromatin remodelers. Mutat Res *618*, 30-40.

131. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal *17*, 10.

132. Maury, Y., Come, J., Piskorowski, R.A., Salah-Mohellibi, N., Chevaleyre, V., Peschanski, M., Martinat, C., and Nedelec, S. (2015). Combinatorial analysis of developmental cues

efficiently converts human pluripotent stem cells into multiple neuronal subtypes. Nat Biotechnol *33*, 89-96.

133. Melki, J., Abdelhak, S., Sheth, P., Bachelot, M.F., Burlet, P., Marcadet, A., Aicardi, J., Barois, A., Carriere, J.P., Fardeau, M.*, et al.* (1990). Gene for chronic proximal spinal muscular atrophies maps to chromosome 5q. Nature *344*, 767-768.

134. Mercuri, E., Finkel, R.S., Muntoni, F., Wirth, B., Montes, J., Main, M., Mazzone, E.S., Vitale, M., Snyder, B., Quijano-Roy, S.*, et al.* (2018). Diagnosis and management of spinal muscular atrophy: Part 1: Recommendations for diagnosis, rehabilitation, orthopedic and nutritional care. Neuromuscul Disord *28*, 103-115.

135. Michalet, X., Ekong, R., Fougerousse, F., Rousseaux, S., Schurra, C., Hornigold, N., van Slegtenhorst, M., Wolfe, J., Povey, S., Beckmann, J.S.*, et al.* (1997). Dynamic molecular combing: stretching the whole human genome for high-resolution studies. Science *277*, 1518-1523.

136. Miller, S., Dykes, D., and Polesky, H. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. Nucl Acid Res, 1216:1215.

137. Mohd-Sarip, A., Teeuwssen, M., Bot, A.G., De Herdt, M.J., Willems, S.M., Baatenburg de Jong, R.J., Looijenga, L.H.J., Zatreanu, D., Bezstarosti, K., van Riet, J.*, et al.* (2017). DOC1-Dependent Recruitment of NURD Reveals Antagonism with SWI/SNF during Epithelial-Mesenchymal Transition in Oral Cancer Cells. Cell Rep *20*, 61-75.

138. Muller, G., Michel, A., and Altenburg, E. (1998). COMP (cartilage oligomeric matrix protein) is synthesized in ligament, tendon, meniscus, and articular cartilage. Connective tissue research *39*, 233-244.

139. Mullis KB, F.F., Scharf SJ, Saiki RK, Horn GT, Ehrlich HA (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harbor Symposium, Quant Biol, 263-273.

140. Mutsaers, C.A., Lamont, D.J., Hunter, G., Wishart, T.M., and Gillingwater, T.H. (2013). Label-free proteomics identifies Calreticulin and GRP75/Mortalin as peripherally accessible protein biomarkers for spinal muscular atrophy. Genome medicine *5*, 95.

141. Neugebauer, J., Heilig, J., Hosseinibarkooie, S., Ross, B.C., Mendoza-Ferreira, N., Nolte, F., Peters, M., Holker, I., Hupperich, K., Tschanz, T.*, et al.* (2018). Plastin 3 influences bone homeostasis through regulation of osteoclast activity. Hum Mol Genet *27*, 4249-4262.

142. Newton, G., Weremowicz, S., Morton, C.C., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., and Lawler, J. (1994). Characterization of human and mouse cartilage oligomeric matrix protein. Genomics *24*, 435-439.

143. Nfonsam, V.N., Nfonsam, L.E., Chen, D., Omesiete, P.N., Cruz, A., Runyan, R.B., and Jandova, J. (2019). COMP Gene Coexpresses With EMT Genes and Is Associated With Poor Survival in Colon Cancer Patients. J Surg Res *233*, 297-303.

144. Nguyen, K., Walrafen, P., Bernard, R., Attarian, S., Chaix, C., Vovan, C., Renard, E., Dufrane, N., Pouget, J., Vannier, A.*, et al.* (2011). Molecular combing reveals allelic combinations in facioscapulohumeral dystrophy. Ann Neurol *70*, 627-633.

145. Ning, Y., Gerger, A., Zhang, W., Hanna, D.L., Yang, D., Winder, T., Wakatsuki, T., Labonte, M.J., Stintzing, S., Volz, N.*, et al.* (2014). Plastin polymorphisms predict gender- and stage-specific colon cancer recurrence after adjuvant chemotherapy. Mol Cancer Ther *13*, 528-539.

146. Novillo, A., Fernandez-Santander, A., Gaibar, M., Galan, M., Romero-Lorca, A., El Abdellaoui-Soussi, F., and Gomez-Del Arco, P. (2021). Role of Chromodomain-Helicase-DNA-Binding Protein 4 (CHD4) in Breast Cancer. Front Oncol *11*, 633233.

147. Olkin, I. (1960). Contributions to probability and statistics: essays in honor of Harold Hotelling (Stanford University Press).

148. Oprea, G.E., Krober, S., McWhorter, M.L., Rossoll, W., Muller, S., Krawczak, M., Bassell, G.J., Beattie, C.E., and Wirth, B. (2008). Plastin 3 is a protective modifier of autosomal recessive spinal muscular atrophy. Science *320*, 524-527.

149. Otsuki, N., Arakawa, R., Kaneko, K., Aoki, R., Arakawa, M., and Saito, K. (2018). A new biomarker candidate for spinal muscular atrophy: Identification of a peripheral blood cell population capable of monitoring the level of survival motor neuron protein. PLoS One *13*, e0201764.

150. Pantel, J.T., Hajjir, N., Danyel, M., Elsner, J., Abad-Perez, A.T., Hansen, P., Mundlos, S., Spielmann, M., Horn, D., Ott, C.E*., et al.* (2020). Efficiency of Computer-Aided Facial Phenotyping (DeepGestalt) in Individuals With and Without a Genetic Syndrome: Diagnostic Accuracy Study. J Med Internet Res *22*, e19263.

151. Pollard, T.D., and Cooper, J.A. (2009). Actin, a central player in cell shape and movement. Science *326*, 1208-1212.

152. Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S*., et al.* (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665-1680.

153. Riessland, M., Ackermann, B., Forster, A., Jakubik, M., Hauke, J., Garbes, L., Fritzsche, I., Mende, Y., Blumcke, I., Hahnen, E*., et al.* (2010). SAHA ameliorates the SMA phenotype in two mouse models for spinal muscular atrophy. Hum Mol Genet *19*, 1492-1506.

154. Riessland, M., Kaczmarek, A., Schneider, S., Swoboda, K.J., Lohr, H., Bradler, C., Grysko, V., Dimitriadi, M., Hosseinibarkooie, S., Torres-Benito, L*., et al.* (2017). Neurocalcin Delta Suppression Protects against Spinal Muscular Atrophy in Humans and across Species by Restoring Impaired Endocytosis. Am J Hum Genet *100*, 297-315.

155. Rigo, F., Chun, S.J., Norris, D.A., Hung, G., Lee, S., Matson, J., Fey, R.A., Gaus, H., Hua, Y., Grundy, J.S*., et al.* (2014). Pharmacology of a central nervous system delivered 2'-O-methoxyethyl-modified survival of motor neuron splicing oligonucleotide in mice and non-human primates. J Pharmacol Exp Ther.

156. Rossoll, W., Jablonka, S., Andreassi, C., Kroning, A.K., Karle, K., Monani, U.R., and Sendtner, M. (2003). Smn, the spinal muscular atrophy-determining gene product, modulates axon growth and localization of beta-actin mRNA in growth cones of motoneurons. J Cell Biol *163*, 801-812.

157. Samstag, Y., and Klemke, M. (2007). Ectopic expression of L-plastin in human tumor cells: diagnostic and therapeutic implications. Adv Enzyme Regul *47*, 118-126.

158. Schrank, B., Gotz, R., Gunnersen, J.M., Ure, J.M., Toyka, K.V., Smith, A.G., and Sendtner, M. (1997). Inactivation of the survival motor neuron gene, a candidate gene for human spinal muscular atrophy, leads to massive cell death in early mouse embryos. Proc Natl Acad Sci U S A *94*, 9920-9925.

159. Schwebach, C.L., Kudryashova, E., Zheng, W., Orchard, M., Smith, H., Runyan, L.A., Egelman, E.H., and Kudryashov, D.S. (2020). Osteogenesis imperfecta mutations in plastin 3 lead to impaired calcium regulation of actin bundling. Bone Res *8*, 21.

160. Scoto, M., Finkel, R.S., Mercuri, E., and Muntoni, F. (2017). Therapeutic approaches for spinal muscular atrophy (SMA). Gene Ther *24*, 514-519.

161. Shanmugarajan, S., Tsuruga, E., Swoboda, K.J., Maria, B.L., Ries, W.L., and Reddy, S.V. (2009). Bone loss in survival motor neuron (Smn(-/-) SMN2) genetic mouse model of spinal muscular atrophy. J Pathol *219*, 52-60.

162. Shapiro, S.S., and Wilk, M.B. (1965). An Analysis of Variance Test for Normality (Complete Samples). Biometrika *52*, 591-&.

163. Shimbo, T., Du, Y., Grimm, S.A., Dhasarathy, A., Mav, D., Shah, R.R., Shi, H., and Wade, P.A. (2013). MBD3 localizes at promoters, gene bodies and enhancers of active genes. PLoS Genet *9*, e1004028.

164. Shinomiya, H. (2012). Plastin family of actin-bundling proteins: its functions in leukocytes, neurons, intestines, and cancer. International journal of cell biology *2012*.

165. Spearman, C. (1904). nthe proof and measurement of association between two things, oamerican J (Psychol).

166. Srinivas, P.R., Wagner, A.S., Reddy, L.V., Deutsch, D.D., Leon, M.A., Goustin, A.S., and Grunberger, G. (1993). Serum alpha 2-HS-glycoprotein is an inhibitor of the human insulin receptor at the tyrosine kinase level. Mol Endocrinol *7*, 1445-1455.

167. Stenvinkel, P., Wang, K., Qureshi, A.R., Axelsson, J., Pecoits-Filho, R., Gao, P., Barany, P., Lindholm, B., Jogestrand, T., Heimburger, O.*, et al.* (2005). Low fetuin-A levels are associated with cardiovascular death: Impact of variations in the gene encoding fetuin. Kidney Int *67*, 2383-2392.

168. Student (1908). The Probable Error of a Mean. Biometrika *6*, 1-25.

169. Sugimachi, K., Yokobori, T., Iinuma, H., Ueda, M., Ueo, H., Shinden, Y., Eguchi, H., Sudo, T., Suzuki, A., Maehara, Y.*, et al.* (2014). Aberrant expression of plastin-3 via copy number gain induces the epithelial-mesenchymal transition in circulating colorectal cancer cells. Annals of surgical oncology *21*, 3680-3690.

170. Sumner, C.J., Kolb, S.J., Harmison, G.G., Jeffries, N.O., Schadt, K., Finkel, R.S., Dreyfuss, G., and Fischbeck, K.H. (2006). SMN mRNA and protein levels in peripheral blood. Biomarkers for SMA clinical trials. Neurology.

171. Szunyogova, E., Zhou, H., Maxwell, G.K., Powis, R.A., Francesco, M., Gillingwater, T.H., and Parson, S.H. (2016). Survival Motor Neuron (SMN) protein is required for normal mouse liver development. Scientific reports *6*, 34635.

172. Tang, N., Gibson, H., Germeroth, T., Porcu, P., Lim, H.W., and Wong, H.K. (2010). T-plastin (PLS3) gene expression differentiates Sezary syndrome from mycosis fungoides and inflammatory skin diseases and can serve as a biomarker to monitor disease progression. Br J Dermatol *162*, 463-466.

173. Thiery, J.P. (2002). Epithelial-mesenchymal transitions in tumour progression. Nat Rev Cancer *2*, 442-454.

174. Tiziano, F.D., Pinto, A.M., Fiori, S., Lomastro, R., Messina, S., Bruno, C., Pini, A., Pane, M., D'Amico, A., Ghezzo, A.*, et al.* (2010). SMN transcript levels in leukocytes of SMA patients determined by absolute real-time PCR. Eur J Hum Genet *18*, 52-58.

175. Tong, J.K., Hassig, C.A., Schnitzler, G.R., Kingston, R.E., and Schreiber, S.L. (1998). Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex. Nature *395*, 917-921.

176. Totzeck, A., Stolte, B., Kizina, K., Bolz, S., Schlag, M., Thimm, A., Kleinschnitz, C., and Hagenacker, T. (2019). Neurofilament Heavy Chain and Tau Protein Are Not Elevated in Cerebrospinal Fluid of Adult Patients with Spinal Muscular Atrophy during Loading with Nusinersen. Int J Mol Sci *20*.

177. Tremblay, D.C., Moseley, S., and Chadwick, B.P. (2011). Variation in array size, monomer composition and expression of the macrosatellite DXZ4. PLoS One *6*, e18969.

178. Tremblay, M., Sanchez-Ferras, O., and Bouchard, M. (2018). GATA transcription factors in development and disease. Development *145*.

179. Tsolis, K.C., Bei, E.S., Papathanasiou, I., Kostopoulou, F., Gkretsi, V., Kalantzaki, K., Malizos, K., Zervakis, M., Tsezou, A., and Economou, A. (2015). Comparative proteomic analysis of hypertrophic chondrocytes in osteoarthritis. Clin Proteomics *12*, 12.

180. Tukey, J.W. (1977). Exploratory Data Analysis (Addison-Wesley).

181. Ueo, H., Sugimachi, K., Gorges, T.M., Bartkowiak, K., Yokobori, T., Muller, V., Shinden, Y., Ueda, M., Ueo, H., Mori, M.*, et al.* (2015). Circulating tumour cell-derived plastin3 is a novel marker for predicting long-term prognosis in patients with breast cancer. Br J Cancer *112*, 1519-1526.

182. van der Leeuw, J., Beulens, J.W., van Dieren, S., Schalkwijk, C.G., Glatz, J.F., Hofker, M.H., Verschuren, W.M., Boer, J.M., van der Graaf, Y., Visseren, F.L.*, et al.* (2016). Novel Biomarkers to Improve the Prediction of Cardiovascular Event Risk in Type 2 Diabetes Mellitus. J Am Heart Assoc *5*.

183. van Dijk, F.S., Zillikens, M.C., Micha, D., Riessland, M., Marcelis, C.L., de Die-Smulders, C.E., Milbradt, J., Franken, A.A., Harsevoort, A.J., Lichtenbelt, K.D.*, et al.* (2013). PLS3 mutations in X-linked osteoporosis with fractures. N Engl J Med *369*, 1529-1536.

184. van Doorn, R., Dijkman, R., Vermeer, M.H., Out-Luiting, J.J., van der Raaij-Helmer, E.M., Willemze, R., and Tensen, C.P. (2004). Aberrant expression of the tyrosine kinase receptor EphA4 and the transcription factor twist in Sezary syndrome identified by gene expression analysis. Cancer Res *64*, 5578-5586.

185. Velthaus, A., Cornils, K., Hennigs, J.K., Grub, S., Stamm, H., Wicklein, D., Bokemeyer, C., Heuser, M., Windhorst, S., Fiedler, W.*, et al.* (2019). The Actin Binding Protein Plastin-3 Is Involved in the Pathogenesis of Acute Myeloid Leukemia. Cancers (Basel) *11*.

186. Viswambharan, V., Thanseem, I., Vasu, M.M., Poovathinal, S.A., and Anitha, A. (2017). miRNAs as biomarkers of neurodegenerative disorders. Biomark Med *11*, 151-167.

187. Wade, P.A., Jones, P.L., Vermaak, D., and Wolffe, A.P. (1998). A multiple subunit Mi-2 histone deacetylase from Xenopus laevis cofractionates with an associated Snf2 superfamily ATPase. Curr Biol *8*, 843-846.

188. Wadman, R.I., Stam, M., Jansen, M.D., van der Weegen, Y., Wijngaarde, C.A., Harschnitz, O., Sodaar, P., Braun, K.P., Dooijes, D., Lemmink, H.H.*, et al.* (2016). A Comparative Study of SMN Protein and mRNA in Blood and Fibroblasts in Patients with Spinal Muscular Atrophy and Healthy Controls. PLoS One *11*, e0167087.

189. Wainer Katsir, K., and Linial, M. (2019). Human genes escaping X-inactivation revealed by single cell expression data. BMC Genomics *20*, 201.

190. Walsh, M.B., Janzen, E., Wingrove, E., Hosseinibarkooie, S., Muela, N.R., Davidow, L., Dimitriadi, M., Norabuena, E.M., Rubin, L.L., Wirth, B.*, et al.* (2020). Genetic modifiers ameliorate endocytic and neuromuscular defects in a model of spinal muscular atrophy. BMC Biol *18*, 127.

191. Wang, L., Zheng, J., Du, Y., Huang, Y., Li, J., Liu, B., Liu, C.J., Zhu, Y., Gao, Y., Xu, Q.*, et al.* (2010). Cartilage oligomeric matrix protein maintains the contractile phenotype of vascular smooth muscle cells by interacting with alpha(7)beta(1) integrin. Circ Res *106*, 514-525.

192. Wang, T.J. (2011). Assessing the role of circulating, genetic, and imaging biomarkers in cardiovascular risk prediction. Circulation *123*, 551-565.

193. Wang, Y., Chen, Y., Bao, L., Zhang, B., Wang, J.E., Kumar, A., Xing, C., Wang, Y., and Luo, W. (2020). CHD4 Promotes Breast Cancer Progression as a Coactivator of Hypoxia-Inducible Factors. Cancer Res *80*, 3880-3891.

194. Wang, Y., Su, M., Zhou, L.L., Tu, P., Zhang, X., Jiang, X., and Zhou, Y. (2011). Deficiency of SATB1 expression in Sezary cells causes apoptosis resistance by regulating FasL/CD95L transcription. Blood *117*, 3826-3835.

195. Wasserman, H.M., Hornung, L.N., Stenger, P.J., Rutter, M.M., Wong, B.L., Rybalsky, I., Khoury, J.C., and Kalkwarf, H.J. (2017). Low bone mineral density and fractures are highly prevalent in pediatric patients with spinal muscular atrophy regardless of disease severity. Neuromuscul Disord *27*, 331-337.

196. Welby, E., Rehborg, R.J., Harmelink, M., and Ebert, A.D. (2022). Assessment of cerebral spinal fluid biomarkers and microRNA-mediated disease mechanisms in spinal muscular atrophy patient samples. Hum Mol Genet *31*, 1830-1843.

197. Wewer, U.M., Ibaraki, K., Schjorring, P., Durkin, M.E., Young, M.F., and Albrechtsen, R. (1994). A potential role for tetranectin in mineralization during osteogenesis. J Cell Biol *127*, 1767-1775.

198. Wilcoxon, F. (1945). Individual Comparisons by Ranking Methods. Biometrics Bulletin *1*, 80-83.

199. Wilczewski, C.M., Hepperla, A.J., Shimbo, T., Wasson, L., Robbe, Z.L., Davis, I.J., Wade, P.A., and Conlon, F.L. (2018). CHD4 and the NuRD complex directly control cardiac sarcomere formation. Proc Natl Acad Sci U S A *115*, 6727-6732.

200. Wirth, B. (2021). Spinal Muscular Atrophy: In the Challenge Lies a Solution. Trends Neurosci *44*, 306-322.

201. Wirth, B., Brichta, L., Schrank, B., Lochmuller, H., Blick, S., Baasner, A., and Heller, R. (2006). Mildly affected patients with spinal muscular atrophy are partially protected by an increased SMN2 copy number. Hum Genet *119*, 422-428.

202. Wolff, L., Strathmann, E.A., Muller, I., Mahlich, D., Veltman, C., Niehoff, A., and Wirth, B. (2021). Plastin 3 in health and disease: a matter of balance. Cell Mol Life Sci *78*, 5275-5301.

203. Wong, H.K., Gibson, H., Hake, T., Geyer, S., Frederickson, J., Marcucci, G., Caligiuri, M.A., Porcu, P., and Mishra, A. (2015). Promoter-Specific Hypomethylation Is Associated with Overexpression of PLS3, GATA6, and TWIST1 in the Sezary Syndrome. J Invest Dermatol *135*, 2084-2092.

204. Woodage, T., Basrai, M.A., Baxevanis, A.D., Hieter, P., and Collins, F.S. (1997). Characterization of the CHD family of proteins. Proc Natl Acad Sci U S A *94*, 11472-11477.

205. Wu, X.S., Lee, S.H., Sheng, J., Zhang, Z., Zhao, W.D., Wang, D., Jin, Y., Charnay, P., Ervasti, J.M., and Wu, L.G. (2016). Actin Is Crucial for All Kinetically Distinguishable Forms of Endocytosis at Synapses. Neuron *92*, 1020-1035.

206. Xin, Z., Li, D., Mao, F., Du, Y., Wang, X., Xu, P., Li, Z., Qian, J., and Yao, J. (2020). PLS3 predicts poor prognosis in pancreatic cancer and promotes cancer cell proliferation via PI3K/AKT signaling. Journal of cellular physiology *235*, 8416-8423.

207. Xiong, F., Wu, G.H., Wang, B., and Chen, Y.J. (2021). Plastin-3 is a diagnostic and prognostic marker for pancreatic adenocarcinoma and distinguishes from diffuse large B-cell lymphoma. Cancer Cell Int *21*, 411.

208. Xu, N., Liu, F., Wu, S., Ye, M., Ge, H., Zhang, M., Song, Y., Tong, L., Zhou, J., and Bai, C. (2020). CHD4 mediates proliferation and migration of non-small cell lung cancer via the RhoA/ROCK pathway by regulating PHF5A. BMC cancer *20*, 262.

209. Xue, Y., Wong, J., Moreno, G.T., Young, M.K., Cote, J., and Wang, W. (1998). NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. Mol Cell *2*, 851-861.

210. Yamada, T., Yang, Y., Hemberg, M., Yoshida, T., Cho, H.Y., Murphy, J.P., Fioravante, D., Regehr, W.G., Gygi, S.P., Georgopoulos, K.*, et al.* (2014). Promoter decommissioning by the NuRD chromatin remodeling complex triggers synaptic connectivity in the mammalian brain. Neuron *83*, 122-134.

211. Yang, F., Babak, T., Shendure, J., and Disteche, C.M. (2010). Global survey of escape from X inactivation by RNA-sequencing in mouse. Genome Res *20*, 614-622.

212. Yang, Y., Yamada, T., Hill, K.K., Hemberg, M., Reddy, N.C., Cho, H.Y., Guthrie, A.N., Oldenborg, A., Heiney, S.A., Ohmae, S.*, et al.* (2016). Chromatin remodeling inactivates activity genes and regulates neural coding. Science *353*, 300-305.

213. Yokobori, T., Iinuma, H., Shimamura, T., Imoto, S., Sugimachi, K., Ishii, H., Iwatsuki, M., Ota, D., Ohkuma, M., Iwaya, T.*, et al.* (2013). Plastin3 is a novel marker for circulating tumor cells undergoing the epithelial-mesenchymal transition and is associated with colorectal cancer prognosis. Cancer Res *73*, 2059-2069.

214. Zachou, K., Gabeta, S., Gatselis, N.K., Norman, G.L., and Dalekos, G.N. (2017). Cartilage oligomeric matrix protein on the spot for liver fibrosis evaluation: Too early or too late? Eur J Intern Med *43*, e48-e49.

215. Zaharieva, I.T., Scoto, M., Aragon-Gawinska, K., Ridout, D., Doreste, B., Servais, L., Muntoni, F., and Zhou, H. (2022). Response of plasma microRNAs to nusinersen treatment in patients with SMA. Ann Clin Transl Neurol *9*, 1011-1026.

216. Zaworski, P., von Herrmann, K.M., Taylor, S., Sunshine, S.S., McCarthy, K., Risher, N., Newcomb, T., Weetall, M., Prior, T.W., Swoboda, K.J.*, et al.* (2016). SMN Protein Can Be Reliably Measured in Whole Blood with an Electrochemiluminescence (ECL) Immunoassay: Implications for Clinical Trials. PLoS One *11*, e0150640.

217. Zeng, Q., Li, L., Feng, Z., Luo, L., Xiong, J., Jie, Z., Cao, Y., and Li, Z. (2021). LCP1 is a prognostic biomarker correlated with immune infiltrates in gastric cancer. Cancer Biomark *30*, 105-125.

218. Zhang, S., Plummer, D., Lu, L., Cui, J., Xu, W., Wang, M., Liu, X., Prabhakar, N., Shrinet, J., Srinivasan, D.*, et al.* (2022). DeepLoop robustly maps chromatin interactions from sparse allele-resolved or single-cell Hi-C data at kilobase resolution. Nat Genet *54*, 1013-1025.

219. Zhang, Y., Castillo-Morales, A., Jiang, M., Zhu, Y., Hu, L., Urrutia, A.O., Kong, X., and Hurst, L.D. (2013). Genes that escape X-inactivation in humans have high intraspecific variability in expression, are associated with mental impairment but are not slow evolving. Mol Biol Evol *30*, 2588-2601.

220. Zhang, Y., LeRoy, G., Seelig, H.P., Lane, W.S., and Reinberg, D. (1998). The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. Cell *95*, 279-289.

221. Zheng, T., Chen, B., Yang, L., Hu, X., Zhang, X., Liu, H., and Qin, L. (2017). Association of plasma dipeptidyl peptidase-4 activity with non-alcoholic fatty liver disease in nondiabetic Chinese population. Metabolism *73*, 125-134.

10. Acknowledgements

Hereby, I want to thank my supervisor Brunhilde Wirth for offering this amazing project to me. I appreciated the freedom that I had during my project. Furthermore, I want to thank my additional supervisors Michael Nothnagel and Martin Peifer for their guidance and ideas to improve my project. I also want to thank my reviewers Niels Gehring and Ulrich Baumann.

Especially, I want to thank Irmgard Hölker for her amazing help in the laboratory, good conversations and her amazing contributions to my project. I know you had a hard time with the ChIP experiments. Furthermore, I want to thank all collaborators that contributed to my work and publications. Lastly, I want to thank all current and former members of the AG Wirth, AG Zempel, AG Kye as well as the diagnostics team. You are a great team and this always helped to keep the mood up. I wish all of you a great future.

Last but not least, I want to thank my family and friends who always believed in me and guided my way. Without you, I would not be the man that I am. Most importantly, I want to thank my wife Leonie and my daughter Madita, who inspire me every day to give my very best in anything that I do. Lastly, I want to thank my parents, Helma and Arnfred, which were always there when I needed them.

11. Zusammenfassung

Erster Teil: Die Spinale Muskelatrophie (SMA) ist eine schwere neurodegenerative Erbkrankheit, die durch homozygote Deletion des Gens Survival of Motor Neuron 1 (*SMN1*) ausgelöst wird. *SMN1*-deletierte Individuen besitzen ein oder mehrere Kopien des Gens *SMN2*, von welchem ungenügende Mengen des SMN Proteins transkribiert werden. Plastin 3 (*PLS3*) ist ein protektiver genetischer Modifikator von SMA. Die Mechanismen, die die Expression von *PLS3* steuern sind weitgehend ungeklärt. In dieser Studie wurden Fibroblasten, Lymphoblasten und spinale Motorneuronen untersucht und Mechanismen entschlüsselt, die die Expression von *PLS3* steuern. In spinalen Motorneuronen wurde eine Verdopplung der Genexpression von *PLS3* in weiblichen asymptomatischen Individuen im Vergleich mit ihren männlichen Geschwistern festgestellt, die sich mit dem Entkommen des Gens aus der X-chromosomalen Inaktivierung erklären lässt. Der *PLS3* Genlokus befindet sich auf dem X-Chromosom nur etwa 500 kb von dem Makrosateliten *DXZ4* entfernt. Dieser zeichnet sich durch eine variable Kopienzahl aus und ist essentiell für die X-chromosomale Inaktivierung. Mittels Molekularem Combing ermittelten wir die genaue Kopienzahl von *DXZ4* und fanden einen linearen Zusammenhang mit der Expression von *PLS3* in Lymphoblastoiden. Darüber hinaus konnten wir das Protein Chromodomain Helicase DNA binding protein 4 (CHD4) als epigenetischen Regulator von *PLS3* identifizieren. Mittels siRNA-induziertem Gen Knock-Down, Überexpressionsstudien, Chromatinimmunopräzipitation und Promoter-Luziferase Assays konnte dieser Zusammenhang validiert werden. Unsere Ergebnisse könnten hilfreich sein, um die Expression von *PLS3* auch im Zusammenhang mit weiteren Erkrankungen zu verstehen.

Zweiter Teil: Spinraza ist ein *SMN* antisense Oligonukleotid, welches in der Therapie von SMA eingesetzt wird. Die Möglichkeit den Fortschritt der Behandlung mittels Biomarkern zu verfolgen ist ein wichtiges Anliegen. Die BforSMA-Studie identifizierte mehr als 200 Biomarker, die mit SMA assoziiert sind. Wir haben die Konzentration von SMN und sechs weiteren Biomarkern (Cartilage Oligomeric Matrix Protein (COMP), Dipeptidyl Peptidase 4 (DPP4), Tetranectin (C-type Lectin Family 3 Member B, CLEC3B), Osteopontin (Secreted Phosphoprotein 1, SPP1), Vitronectin (VTN) und Fetuin A (Alpha 2-HS Glycoprotein, AHSG)) in einem intermediären SMA-Mausmodel untersucht. Dieses Mausmodell wurde präsymptomatisch mit geringdosierten SMN-

Antisenseoligonukleotiden behandelt. Außerdem wurde die Auswirkung von *PLS3*-Überexpression mittels eines humanen Transgens auf die Plasma Biomarker untersucht. Lediglich COMP, DPP4 und SPP1 zeigten eine Korrelation mit dem SMA Phänotyp. Die Überexpression von *PLS3* zeigte keine Auswirkung auf SMN und die übrigen sechs Biomarker.

12. Erklärung zur Dissertation

Erklärung zur Dissertation gemäß der Promotionsordnung vom 02. Februar 2006 mit den Änderungsordnungen vom 10. Mai 2012, 16. Januar 2013 und 21. Februar 2014.

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit − einschließlich Tabellen, Karten und Abbildungen −, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie − abgesehen von unten angegebenen Teilpublikationen − noch nicht veröffentlicht worden ist, sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Brunhilde Wirth betreut worden.

Teilpublikationen:

2022: Scientific publication as author
Strathmann EA, Hölker I, Tschernoster N, Hosseinibarkooie S, Come J, Martinat C, Altmüller J, Brunhilde Wirth B Epigenetic regulation of plastin 3 expression by the macrosatellite DXZ4 and the transcriptional regulator CHD4
Under revision at AJHG

2021: Review as shared author
Wolff L, **Strathmann EA**, Müller I, Mählich D, Veltmann C, Niehoff A, Wirth B
Plastin 3 in health and disease: a matter of balance Cell Mol Life Sci. 2021 Jul;78(13):5275-5301.doi: 10.1007/s00018-021-03843-5. Epub 2021 May 23.

2018: Scientific publication as author
Strathmann EA, Peters M, Hosseinibarkooie S, Rigo FW, Bennett CF, Zaworski PG, Chen KS, Nothnagel M, Wirth B
Evaluation of potential effects of Plastin 3 overexpression and low-dose SMN-antisense oligonucleotides on putative biomarkers in spinal muscular atrophy mice. PLoS One. 2018 Sep 6;13(9):e0203398. doi: 10.1371/journal.pone.0203398. eCollection 2018.

19.08.22 Strathmann

Datum / Unterschrift