Dissertation

submitted to the

Combined Faculty of Natural Sciences and Mathematics

of the Ruperto Carola University Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Presented by

M.Sc. Tanja Mederer

born in: Darmstadt, Germany

Oral examination: 8.10.2019

Identification and characterization of novel candidate genes for *Hirschsprung's disease* –

a developmental disorder of the enteric nervous system

Referees: Prof. Dr. Gudrun Rappold

Prof. Dr. Beate Niesler

ABSTRACT

The neurodevelopmental disorder Hirschsprung's disease (HSCR) represents the most common cause for congenital obstruction and is characterized by a lack of enteric neurons (aganglionosis) in distinct segments of the colon. This aganglionosis is caused by dysfunctions in the neural crest cell (NCC) population which is responsible for enteric nervous system (ENS) generation during embryogenesis. Specifically, either proliferation, migration, differentiation or cell survival of NCC-derived progenitor cells is impaired. The main symptom of this disorder is represented by a megacolon formation. Patients are routinely treated by surgical resections of the aganglionic segment, but gastrointestinal impairments may persist in a fraction of patients even life-long. HSCR is classified as rare and multifactorial disorder. Up to now more than 20 genes are classified as validated disease-causing loci, including the major susceptibility locus *RET*, but many more genetic factors have been implicated in the pathoaetiology. However, in the majority of patients the genetic disease causes are still unknown. Next-generation sequencing technologies provide the possibility to rapidly uncover the individual's genetic architecture. Nevertheless, dissecting the genetic findings of importance and correlating them with the pathomechanisms is a major challenge especially in complex diseases as HSCR. This study aimed to establish a complementary research approach for identification and characterization of novel HSCR candidate genes. By taking genetic, bioinformatics, molecular and functional data into account, better insights into the molecular pathogenesis of HSCR should be gained.

In this project, two sporadic long-segment HSCR cases were analysed by whole exome sequencing in a trio-based setup and by genotyping of non-coding risk single nucleotide polymorphisms. In both patients, bioinformatic analyses of exome-wide sequencing data led to the identification of rare structural (copy number variations (CNVs)) and single nucleotide variants (SNVs). To narrow down the list of HSCR candidate genes, rare SNVs were further filtered. Finally, four candidate genes (*ATP7A, SREBF1, ABCD1* and *PIAS2*) which were so far not reported in the context of HSCR, were selected for detailed investigations. Extensive mRNA and protein expression analyses confirmed the expression of these candidates in relevant murine gastrointestinal tissues of different developmental stages and thereby validated their putative relevance for the HSCR aetiology. Moreover, additional HSCR patients carrying rare variants in *SREBF1* and *PIAS2* were identified. To further assess functionally the neuronal specific role of the candidates, the CRISPR/Cas9 technology was applied in a human neuroblastoma cell line. Gene-specific *knockout* (*KO*) cell clones were generated for three candidate genes and the major HSCR susceptibility locus *RET,* while genome editing was not successful for *PIAS2*. *KO* clones were investigated on morphological and functional level by a comparison to a *mock control* clone. Comparative analyses revealed variable differences for the individual gene-specific *KO* clones in the differentiation behaviour, proliferation and migration capacity as well as in cell survival during neuronal differentiation. To evaluate all findings of this complementary project, a HSCR risk scoring system was applied. According to the gained risk scores, all four selected HSCR candidates could be classified as relevant for the development of HSCR. Like this, the suitability of the presented research approach for identification and characterization of novel HSCR candidates was validated. It is envisioned to apply the established study pipeline to primary ENS-like model systems to confirm the findings of this project. Moreover, these analyses could help to dissect the candidate gene's relevance for HSCR in detail and gain insights into affected molecular pathways.

ZUSAMMENFASSUNG

Die neurobiologische Entwicklungsstörung Morbus Hirschsprung (MH) repräsentiert die häufigste Ursache für angeborene Darmobstruktion und ist durch das Fehlen von enterischen Neuronen (Aganglionose) in spezifischen Abschnitten des Kolons charakterisiert. Diese Aganglionose resultiert aus Dysfunktionen in den Neuralleistenzellen (NLZ), welche für die Bildung des enterischen Nervensystems (ENS) im Rahmen der Embryogenese verantwortlich sind. Im Speziellen sind die NLZ-abgeleiteten Vorläuferzellen entweder in ihrer Proliferation, der Migration, der Differenzierung oder dem Zellüberleben gestört. Das Hauptsymptom von MH ist die Bildung eines Megakolons. Die standardmäßige Behandlung von MH-Patienten liegt in der operativen Entfernung des aganglionären Darmsegmentes, wobei funktionelle Darmbeschwerden in einem Teil der Patienten auch lebenslang persistieren. Auf genetischer Ebene wird MH als seltene und multifaktorielle Erkrankung klassifiziert. Bis heute sind mehr als 20 Gene, inklusive des bisher beschriebenen Hauptrisikolocus *RET*, als krankheitsauslösend beschrieben und viele weitere Gene wurden mit der Pathogenese in Verbindung gebracht. Dennoch ist die genetische Ursache in der Mehrheit der MH-Patienten bisher ungeklärt. *Next-generation-sequencing* Verfahren ermöglichen die schnelle Entschlüsselung der genetischen Architektur von Individuen. Die Auswahl der relevantesten genetischen Informationen und deren Korrelation mit den Pathomechanismen ist jedoch insbesondere im Zusammenhang mit komplexen Erkrankungen, wie MH, eine große Herausforderung. Zielsetzung dieses Projektes war die Etablierung einer komplementären Studienpipeline zur Identifizierung und Charakterisierung neuer MH-Kandidatengene. Die Betrachtung von genetischen, bioinformatischen, molekularen und funktionellen Daten soll zu einem besseren Verständnis der molekularen MH-Pathomechanismen beitragen.

Im Rahmen dieser Studie wurden zwei sporadische, langstreckige MH-Patienten mittels Trio-basierter Exom-weiter Sequenzierung analysiert und für nicht-codierende Risiko-Polymorphismen genotypisiert. Die bioinformatische Analyse der Exom-weiten Sequenzdaten führte zur Identifizierung von seltenen strukturellen (Kopienzahlvariationen) und Einzelnukleotid-Varianten in beiden Patienten. Weitere Filterschritte wurden genutzt, um aus der Liste an seltenen Einzelnukleotid-Varianten die vielversprechendsten Genloci auszuwählen. Schlussendlich konnten somit vier Kandidatengene *(ATP7A, SREBF1, ABCD1* und *PIAS2)*, die bisher nicht mit MH beschrieben worden waren, für weitere Analysen bestimmt werden. Detaillierte Expressionsanalysen auf mRNA und Proteinebene zeigten, dass alle Kandidaten in murinem gastrointestinalem Gewebe zu relevanten Entwicklungsstadien exprimiert werden. Somit bestätigte sich eine putative Relevanz dieser Gene für die MH-Pathogenese. Zusätzlich konnten weitere MH-Patienten als Träger von seltenen Varianten in *SREBF1* oder *PIAS2* identifiziert werden. Um die neuronale Rolle der Kandidatengene experimentell zu untersuchen, wurde die CRISPR/Cas9 Methode in einer humanen Neuroblastomzelllinie angewandt. Gen-spezifische *Knockout* (*KO*) Zellklone konnten für drei Kandidaten und den bisher beschriebenen MH-Hauptrisikolocus *RET* generiert werden, wobei die Genomeditierung für *PIAS2* nicht erfolgreich war. Die Gen-spezifischen *KO* Klone wurden für morphologische und funktionelle Analysen verwendet und hierbei mit einem *Mock* Kontrollklon verglichen. Diese vergleichenden Analysen zeigten variable Unterschiede in den individuellen *KO* Klonen im Differenzierungs-, Proliferations-, Migrations- und Zellapoptose-Verhalten. Um alle Ergebnisse dieser Studie abschließend bewerten zu können, wurde ein Risiko-Klassifizierungssystem genutzt. Die entsprechenden Risikowerte belegten eine MH-

Relevanz für alle vier Kandidaten. Auf diese Weise konnte die Eignung der komplementären Studienpipeline zur Identifizierung und Charakterisierung neuer MH-Kandidatengene validiert werden. Um die Daten dieser Arbeit zu bestätigen, ist anvisiert, die etablierte Studienpipeline zukünftig auch in einem primären ENS-ähnlichen Zellkulturmodellsystem anzuwenden. Die entsprechenden Analysen könnten dazu beitragen, die Relevanz der Kandidatengene detailliert zu analysieren und somit Einblick in die betroffenen molekularen Signalwege zu erhalten.

PRESENTATIONS AND PUBLICATIONS

Data generated within this project were already/will be presented at the following conferences:

- Retreat of the ”Interdisciplinary center for neurosciences” (IZN) Heidelberg;
  July, 2016, 2018 and 2019, Schöntal (Germany); Poster

- 24th annual meeting of the German Society of Neurogastroenterology and Motility;
  March 10-12, 2017, Berlin (Germany); Talk

- NeuroGASTRO 2019 Biennal Meeting of the European Society of Neurogastroenterology and Motility; September 5-7, 2019, Lisbon (Portugal); Poster

The following publications resulted/will result from this project and additional work:

- Schmitteckert, S., **Mederer, T.**, Niesler, B.; 2016; *„Genetik der Darm-Hirn-Achse am Beispiel des Reizdarmsyndroms“*, NeuroLogisch, Focus: Neurogastroenterologie, neuro 02|2016 08.07.2016.

- #Wohlfarth, C., #Schmitteckert, S., Härtle, J.D., Houghton, L.A., Dweep, H., Fortea, M., Assadi, G., Braun, A., **Mederer, T.**, Pöhner, S., Becker, P.B., Fischer, C., Granzow, M., Mönnikes, H., Mayer, E.A., Sayuk, G., Boeckxstaens, G., Wouters, M.M., Simrén, M., Lindberg, G., Ohlsson, B., Schmidt, P.T., Dlugosz, A., Agreus, L., Andreasson, A., D’Amato, M., Burwinkel, B., Lorenzo, J., Röth, R., Lasitschka F., Vicario, M., Metzger, M., Santos, J., Rappold, G., *Martinez, C., *Niesler, B.; 2017; *“miRNA-16 and miR-103 impact 5-HT4 receptor signalling and correlate with symptom profile in irritable bowel syndrome“*, SciRep, 31;7(1):14680. doi: 10.1038/s41598-017-13982-0., #joined first authorship, *joined senior authorship

- Schmitteckert, S.; **Mederer, T.**; Röth, R.; Günther, P.; Holland-Cunz, S.; Metzger, M.; Schröder-Braunstein, J.; Heidtmann, A.; Scheuerer, J.; Beretta, C. A.; Lasitschka, F.; Rappold, G. A.; Romero, P.; Niesler, B.; 2019; *“Postnatal human enteric neurospheres show a remarkable molecular complexity”*; Neurogastroenterology & Motility

- **Mederer, T.**, Schmitteckert, S., Volz, J., Martinez, C., Röth, R., Thumberger, T., Eckstein, V., Tapia-Laliena, M., Khasanov, R., Schäfer, KH., Wessel, L., Scheuerer, J., Thöni, C., Lasitschka, F., Carstensen, L., Günther, P., Holland-Cunz, S., Hofstra, R., Paramasivam, N., Rappold, G.A., Romero, P., Niesler, B., & International Hirschsprung Disease Consortium. *“Complementary study pipeline to elucidate the complex genetics of Hirschsprung‘s disease”*, in preparation

CONTRIBUTIONS

Experimental data presented in this thesis was partly generated by or in cooperation with other people and Core facilities.

Julia Volz performed the initial candidate gene selection based on a literature and database search as well as a network analysis using filtered whole exome sequencing datasets from the two index cases and their relatives that were followed up in this PhD thesis (Master Thesis Volz, 2014).

The whole exome sequencing data analysis and subsequent bioinformatic filtering analyses were performed by **Dr. Nagarajan Paramasivam** (Theoretical Bioinformatics Division, DKFZ).

Murine embryonic tissue for mRNA expression profiling was dissected together with **Dr. Stefanie Schmitteckert**. Microarray analyses were carried out at the Genomics and Proteomics Core Facility of the DKFZ (**Dr. Melanie Bewerunge-Hudler**). Respective data analysis was performed by **Dr. Carolina de la Torre** from the Center of Medical Research (Medical Faculty Mannheim of the University of Heidelberg).

Preparation, sectioning as well as immunohistochemical staining of human formalin-fixed paraffin embedded specimens were performed by **Dr. Cornelia Thöni** and **Jutta Scheuerer** (research group of Dr. Felix Lasitschka, Institute of Pathology, University Hospital Heidelberg).

Fluorescent-activated cell sorting was carried out at the FACS Core Facility of the University Hospital Heidelberg (**Dr. Volker Eckstein**).

The scoring system for risk evaluation of novel candidate genes was established by **Denise Dawid** (under my supervision) (Bachelor Thesis Dawid, 2019).

LIST OF CONTENTS

ABBREVIATIONS

ABC	ATP binding cassette
ABCD1	*ABC SUBFAMILY MEMBER 1*
ABCD2	ABC SUBFAMILY MEMBER 2
ACh	ACETYLCHOLINE
A-domain	Actuator domain
AGPAT9	*GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE 3*
Artn	Artemin
ASCL1	*ACHAETE-SCUTE FAMILY BHLH TRANSCRIPTION FACTOR 1*
ASD	Autism Spectrum Disorder
ATP	Adenosine triphosphate
ATP7A	*ATPase COPPER TRANSPORTING ALPHA*
b	base
BAX	*BCL2-ASSOCIATED X*
BCL2	*B-CELL LYMPHOMA 2*
Bdnf	Brain-derived neurotrophic factor
BHLH	Basic Helix Loop Helix
BMP	BONE MORPHOGENETIC PROTEIN
bp	base pairs
BrdU	Bromodeoxyuridine
BSA	Bovine Serum Albumin
C	Celcius
CADD	Combined annotatation dependent depletion
CFU	Colony forming units
CLN8	*CLN8 TRANSMEMBRANE ER AND ERGIC PROTEIN*
Da	Dalton
DAB	3,3′-Diaminobenzidine
DNA	Deoxyribonucleic acid
CASP3	*CASPASE3*
CD1	Cluster of differentiation 1
cDNA	complementary DNA
CGH	Comparative genomic hybridization
CM	Culture medium
CM layer	Circular muscle layer
CNS	Central nervous system
CNV	Copy number variation
Col I	Collagen I
Col IV	Collagen IV
comp.	compound
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
C-terminal	Carboxy terminal
d	days
DEPC	Diethyl pyrocarbonate
DKFZ	German Cancer Research Center
DM	Differentiation medium
DMEM	Dulbecco’s modified eagle medium
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid

DNMT3B	*DNA METHYLTRANSFERASE 3 BETA*
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen
E.Col	*Escherichia coli*
ECad	Epithelial Cadherin
ECE1	*ENDOTHELIN CONVERTING ENZYME 1*
Edn1,2, 3	Endothelin 1,2,3
Ednrb	Endothelin receptor type B
EDTA	Ethylenediaminetetraacetic acid
EEF1A2	*EUKARYOTIC TRANSLATION ELONGATION FACTOR 1 ALPHA 2*
EMT	Epithelium-to-mesenchyme transition
ENCDCs	Enteric neural crest derived cells
ENS	Enteric nervous system
ER	Endoplasmic reticulum
EST	Expressed sequence tag
EtOH	Ethanol
EUR	European
ExAC	Exome Aggregation Consortium
ExPASy	Expert Protein Analysis System
Fabp	Fatty-acid-binding protein
FAC	Fluorescent-activated cell
FACS	Fluorescent-activated cell sorting
FBS	Fetal bovine serum
FFPE	Formalin-fixed paraffin-embedded
FGF	FIBROBLAST GROWTH FACTOR
FGIC_C	Human fetal gastrointestinal cells from the colon
FL	full-length
g	gram
GABA	GAMMA AMINOBUTYRIC ACID
GAP43	*GROWTH-ASSOCIATED PROTEIN-43*
GAPDH	*GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE*
GATA2	GATA BINDING PROTEIN 2
Gdnf	Glial-derived neurotrophic factor
GFP	Green fluorescent protein
Gfap	Glial fibrillary acidic protein
Gfrα1	Gdnf family receptor alpha 1
GI	Gastrointestinal
gnomAD	Genome Aggregation Database
GPI	glycosylphosphatidylinositol
GTEx	Genotype Tissue expression
GWAS	Genome wide association study
h	hours
HEK	Human embryonic kidney
HLH	helix loop helix
HMA	heavy metal associated
HPRT1	*HYPOXANTHINE PHOSPHORIBOSYLTRANSFERASE 1*
HSCR	Hirschsprung's disease
HuC/D	Hu-Antigen C/D
IBD	Inflammatory bowel disease
IBF	Interfacultary Biomedical Faculty

IBS	Inflammatory bowel syndrome
IF	Immunofluorescence
IgG	Immunoglobulin G
IHC	Immunohistochemical
indel	insertion/deletion
IPA	Ingenuity pathway analysis
IPANs	Intrinsic primary afferent neurons
iPSC	induced pluripotent stem cell
IPTG	Isopropyl β-D-1-thiogalactopyranoside
k	kilo
Ki67	Antigen Ki-67
KIAA1279	KIF1 BINDING PROTEIN
KO	Knockout
L	liter
L1CAM	*L1 CELL ADHESION MOLECULE*
LB	Lysogeny broth
LDL	LOW-DENSITY LIPOPROTEIN
L-HSCR	long-segment HSCR
LM layer	Longitudinal muscle layer
LMNA	LAMININ A/C
LoF	Loss-of-function
M	mol/L
m	milli or metre
MAF	Minor allele frequency
MAP2	*MICROTUBULE-ASSOCIATED PROTEIN 2*
Mash1	Achaete-Scute Family BHLH Transcription Factor 1
MAT	mature
MEN2A	Multiple endocrine neoplasia type 2A
MeOH	Methanol
MH	Morbus Hirschsprung
MICA	*MHC CLASS I POLYPEPTIDE-RELATED SEQUENCE A*
min	minute
MLPA	Multiplex ligation-dependent probe amplification
MP	Myenteric plexus
mRNA	messenger RNA
MUT	mutated
n	nano
NAV2	NEURON NAVIGATOR 2
NC	Neural crest
NCBI	National Center for Biotechnology Information
NCC	Neural crest cell
N-domain	Nucleotide domain
NES	*NESTIN*
Ngf	Nerve growth factor
NGS	Normal goat serum
NHEJ	Non-homologous end joining
NLS	Nuclear localisation signal
NLZ	Neuralleistenzellen
NMD	Nonsense-mediated decay

nNOS	neuronal NITRIC OXIDE SYNTHASE
Npy	Neuropeptide Y
NRG1,3	*NEUREGULIN 1,3*
Nrtn	Neurturin
Nt3, 4	Neutrophin 3,4
N-terminal	Amino terminal
NTF3	*NEUTROPHIN 3*
NTRK3	*NEUROTROPHIC RECEPTOR TYROSINE KINASE 3*
o/n	overnight
oe	observed/expected
OMIM	Online Mendelian Inheritance in Man
OptiMEM	Opti Minimal Essential Medium
p75	Low-affinity neutrophin receptor p75
P75NTR	*LOW-AFFINITY NEUTROPHIN RECEPTOR P75NTR*
PAM	Protospacer adjacent motif
PBS	Phosphate-buffered saline
PC	Polycarbonate
PCDHGA1/2/3/4, PCDHGB1/2	*PROTO-CADHERIN GAMMA SUBFAMILY A (1/2/3/4), B (1,2)*
PCR	Polymerase chain reaction
P-domain	Phosphorylation domain
PFA	Paraformaldehyde
Pgp9.5	Ubiquitin C-Terminal Hydrolase L1
Phox2b	Paired-like homeobox 2b
PIAS2	*PROTEIN INHIBITOR OF ACTIVATED STAT 2*
PINIT	Pro-Ile-Asn-Ile-Thr
PROK1	*PROKINETICIN 1*
PROKR1,2	*PROKINETICIN RECEPTOR 1,2*
PROVEAN	Protein variation effect analyzer
Pspn	Persephin
PTEN	PHOSPHATASE AND TENSIN HOMOLOG
PVDF	Polyvinylidene fluoride
qRT	quantitative real-time
RA	Retinoic acid
RARB	RETINOIC ACID RECEPTOR BETA
RBP4	RETINOL BINDING PROTEIN 4
RMA	Robust multi-array
Ret	Rearranged during transfection
RIPA	Radio Immuno Precipitation Assay
RNA	Ribonucleic acid
rpm	rounds per minute
rs	RefSNP
RT	room temperature
S100b	S100 Calcium binding Protein b
SAP	Scaffold Attachment Factor A/B acinus and PIAS
SDHA	*SUCCINATE DEHYDROGENASE COMPLEX FLAVOPROTEIN SUBUNIT A*
SDS	Sodium dodecyl sulphate
SEM	Standard error of mean
SEMA3A/C/D	*CLASS 3 SEMAPHORIN A/C/D*
SERPINI 1	*SERPIN FAMILY I MEMBER 1*

sgRNA	single guide RNA
SHH	SONIC HEDGEHOG
S-HSCR	short-segment HSCR
Sma	Smooth muscle actin
SMP	Submucosal plexus
SNP/SNV	Single nucleotide polymorphism/single nucleotide variant
SOB	Super optimal broth
SOC	SOB with catabolite repression
Sox10	SRY Box 10
SPF	specific pathogen free
SREBF1	*STEROL REGULATORY ELEMENT BINDING TRANSCRIPTION FACTOR 1*
SRY	Sex determining region Y
STAT	SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION
STD	Standard deviation
SUFU	*SUPRESSOR OF FUSED HOMOLOG*
SUMO	Small Ubiquitin-Related Modifier
SYP	*SYNAPTOPHYSIN*
TAC1	*TACHYKININ PRECURSOR 1*
Tac1	Substance P
TAE	TRIS-Acetate-EDTA
TAU	*MICROTUBULE-ASSOCIATED PROTEIN TAU*
TBS	TRIS-buffered saline
TBST	TRIS-buffered saline with Tween-20
TBX3	T-BOX TRANSCRIPTION FACTOR 3
TCA	Total colonic aganglionosis
TCF4	*TRANSCRIPTION FACTOR 4*
TEMED	N,N,N′,N′-Tetramethylethylenediamine
TM	transmembrane
TNF	Tumor necrosis factor
Tubb3	Beta III-Tubulin
TUNEL	TdT-mediated dUTP-X nick end labelling
UCHL1	*UBIQUITIN C-TERMINAL HYDROLASE L1*
UCSC	University of California, Santa Cruz
UHRF1BP1L	*UHRF1 BINDING PROTEIN 1 LIKE*
V	Volt
VIP	VASOACTIVE INTESTINAL PEPTIDE
VLCFA	Very long change fatty acid
WES	Whole exome sequencing
WNT	WINGLESS
wog	week of gestation
WT	wildtype
X-ALD	X-linked adrenoleukodystrophy
µ	micro
ZFHX1B	*ZINC FINGER E-BOX BINDING HOMOEOBOX*
ZF-SP-RING	Zinc Finger-Really Interesting New Gene

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

1.1 The enteric nervous system

The enteric nervous system (ENS) resides within the gut wall and is the largest unit of the peripheral nervous system. Its main function is the intrinsic innervation of the gastrointestinal (GI) tract, which regulates the blood flow, gut motility and peristalsis, ion and fluid homeostasis as well as secretion of signalling mediators (Rao and Gershon, 2018). The ENS shares basic concepts with the central nervous system (CNS) such as the composition of signalling molecules and neurotransmitters, but can operate autonomously from its input. Therefore, it has been coined the “second brain” (Furness and Stebbing, 2018; Sasselli *et al.*, 2012). The bidirectional communication between the ENS and the CNS is designated as gut-brain axis (Mayer and Tillisch, 2011).

1.1.1 *Structure and function*

The basic structure of the ENS includes enteric ganglia organized in two major plexuses: corresponding to their location, these are referred to as the outer myenteric (Auerbach’s) plexus situated between the circular and the longitudinal muscle layer, and the inner submucosal (Meissner’s) plexus residing within the submucosa (Rao and Gershon, 2018) (fig. 1 A). During embryonic development the myenteric plexus is generated first and can be found throughout the GI tract, whereas the submucosal plexus derived by a second radial migration front, is present in the small and large intestine only (Heanue and Pachnis, 2007). Enteric ganglia in both plexus are interconnected and are composed of various neuronal subtypes and glia. Besides the complex network of enteric ganglia, various other cell types are required for neuronal circuit integration (Belkind-Gerson *et al.*, 2006). Consequently, tight communication between enteric ganglia and epithelial, endocrine, immune cells as well as with interstitial cells of Cajal and smooth muscle cells is crucial for proper GI function (Heanue and Pachnis, 2007; Rao and Gershon, 2018).

Neuronal subtypes organized in enteric ganglia can be differentiated based on their neurochemical coding and their electrophysiological properties. In principle, neurons are classified as intrinsic primary afferent neurons (IPANs), inter- and motorneurons (Rao and Gershon, 2018). Projections of IPANs spread into intestinal villi and respond to chemical and mechanical stimuli (fig. 1 B). Therefore, IPANs are crucial for the regulation of ENS autonomous functions. The coordinated action of these in concert with afferent and efferent neurons innervating the circular muscle layer orchestrate the movement of the digestive tract. Gap junctions between smooth muscle cells transmit the electrical signal and thus ensure propulsion and mixing of the ingested food via peristalsis. Secretomotor and vasomotor neurons of the submucosal plexus innervate effector cells as enteroendocrine cells and blood vessels, respectively (Benarroch, 2007; Rao and Gershon, 2018). Consequently, secretion of hormones and digestive enzymes by the mucosa and fluid exchange through vasodilation are integrated as GI functions (Belkind-Gerson *et al.*, 2006).

In addition to the electrophysiological characteristics of enteric neurons, neuronal diversity is based on cell type specific neurotransmitter based chemical coding (Heanue and Pachnis, 2007). While most of the myenteric neurons are cholinergic, neuronal NITRIC OXIDE SYNTHASE (nNOS) is classically found in motor- and interneurons. Serotonin serves as neurotransmitter in interneurons as well, while GAMMA AMINOBUTYRIC ACID (GABA) and dopamine are further neuroactive molecules found in the ENS. Secretomotorneurons and IPANs of the submucosal plexus express VASOACTIVE INTESTINAL PEPTIDE (VIP) and ACETYLCHOLINE (ACh). The complex diversity in the neurotransmitter profile of enteric neurons is not fully understood so far (Hao and Young, 2009; Rao and Gershon, 2018; Sasselli *et al.*, 2012).

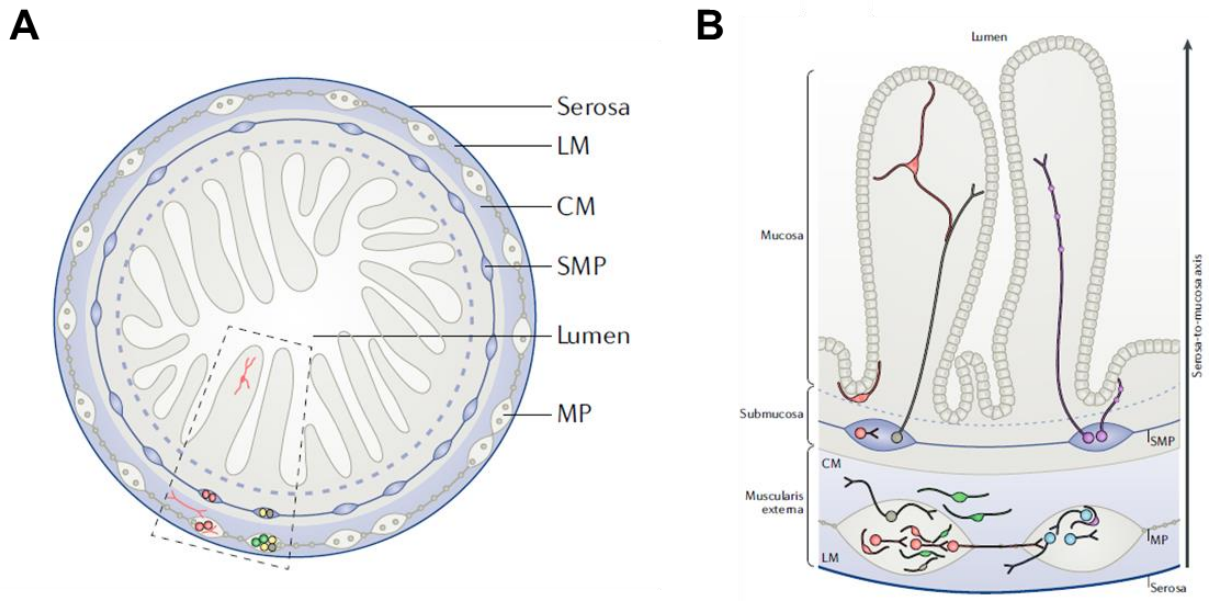


Figure 1: Schematic of the human ENS organization
A: Crosssection of the postnatal gut displaying the organization of the ENS in two major plexus (SMP: submucosal plexus, MP: myenteric plexus). The MP is embedded between the longitudinal and circular muscle layer (LM and CM). The inset region is shown as a blow up in **B:** Enteric ganglia are interconnected and consist of different cell types (blue: neuron, green: glia, red: clonally related neurons and glia). IPANs (purple) project to the intestinal villi. (modified from Rao and Gershon, 2018).

1.1.2 *Development*

The ENS has its origin in early embryonic development, since it arises from neural crest cells (NCCs). This highly proliferative and migratory cell population is specified during neurulation, as illustrated schematically in figure 2 A. The late gastrula stage embryo contains three regions of ectodermal origin: the non-neural ectoderm, the neural plate border and the neural plate (neural ectoderm). Formation of the neural tube is initiated when the neural plate borders rise and form the neural folds. The neural tube is the embryonic precursor of the CNS. Fusion of the neural folds at the midline causes the involution of the neural tube into the embryo. The non-neuronal epidermis develops into the epidermis covering the neural tube. At the dorsal fusion sites of the neural folds, the neural crest is specified (Bhatt *et al.*, 2013; Gammill and Bronner-Fraser, 2003; Gilbert, 2014). After epithelium-to-mesenchyme transition (EMT), NCCs delaminate from this transient structure and migrate to different regions throughout the body,

where they give rise to various cell types and lineages (Bhatt *et al.*, 2013; Mayor and Theveneau, 2013) (fig. 2 A). Amongst others, neuronal and glial cells, melanocytes and endocrine cells as well as cartilage and bone are generated. Corresponding to early phases of neurulation, where BONE MORPHOGENETIC PROTEIN (BMP), WINGLESS (WNT) and FIBROBLAST GROWTH FACTOR (FGF) signalling molecules play central roles in neural plate induction, these signalling pathways are as well involved in NCC specification (Bhatt *et al.*, 2013; Mayor and Theveneau, 2013; Simões-Costa and Bronner, 2015).

NCCs are classified into specific groups according to their position along the anterior/posterior body axis and corresponding to their progenitor potential: in general, cranial, vagal, trunk and sacral NCCs can be differentiated (Simões-Costa and Bronner, 2015). While sacral NCCs (originating posterior to somite 28) contribute only minimally to the later ENS cell population, most enteric ganglia are derived from vagal NCCs. These progenitor cells are specified at the level of the first to seventh somite. Additionally, a small proportion of the most anterior specific trunk NCCs contribute minimally to ENS formation in the foregut. Delamination of the vagal NCC subpopulation is initiated between E8.5-9 in mouse and prior to week 4 gestation in a human embryo (Gilbert, 2014; Goldstein *et al.*, 2013; Heanue and Pachnis, 2007; Lake and Heuckeroth, 2013) (fig. 2 B).

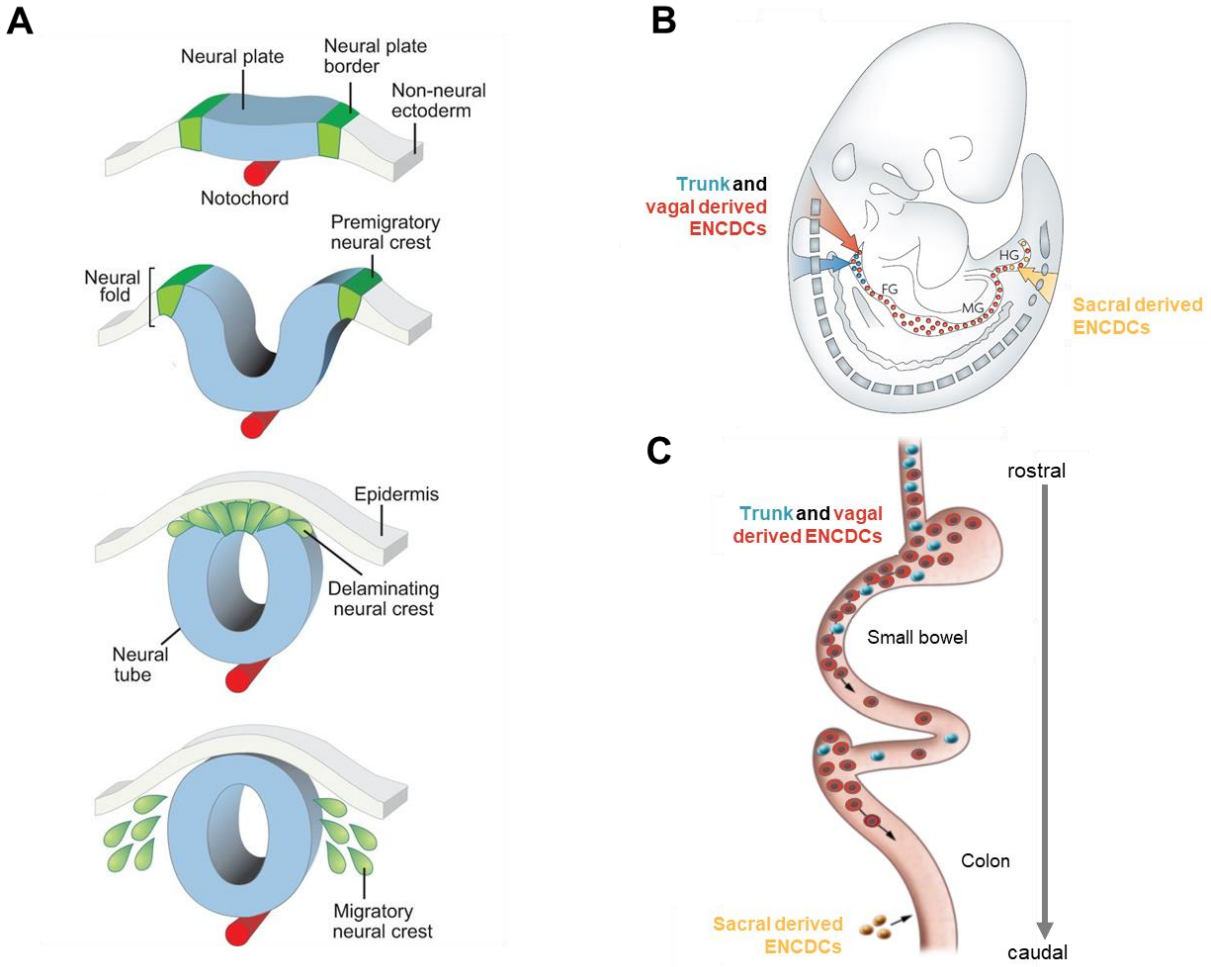


Figure 2: Different phases of ENS development
A: Delamination of NCCs (green) from the neural tube (blue) in the context of neurulation (modified from Simões-Costa and Bronner, 2015). **B:** Different NCC subpopulations (trunk (blue), vagal (red) and sacral (yellow)) invade and colonize the gut (modified from Heanue and Pachnis, 2007). **C:** Gut colonization by trunk (blue) and vagal ENCDCs (red) in rostral to caudal direction (modified from Heuckeroth, 2015).

After delamination, these multipotent progenitor cells migrate ventrally through the somatic mesenchyme to reach the presumptive gut. After invasion of the developing bowel, ENS progenitor cells, which are termed enteric neural crest derived cells (ENCDCs), start to colonize the embryonic gut. At E9.5, vagal derived ENCDCs migrate in rostro-caudal direction within the developing mid- and hindgut. In contrast, sacral derived ENS precursor cells do not enter the gut before E13.5 and migrate orally (fig 2 B and C) (Heanue and Pachnis, 2007; Rao and Gershon, 2018). Gut colonization by vagal ENCDCs is completed around E14 in mouse and after 7 weeks gestation in human embryos. To generate the two interconnected ganglia layers, referred to as myenteric and submucosal plexus, a second radial directed migration wavefront follows the first linear migration wavefront (Lake and Heuckeroth, 2013). Except for gut invasion and colonization, ENCDCs cells are responsible for the generation of the ENS-specific neuronal and glial subtype diversity, which is achieved by respective cell fate commitment. Molecular mechanisms that regulate all these processes are presented in 1.1.3.

Even after ENS formation is completed, stem and progenitor cells persist in both ganglionated plexus and the outer muscle layer. These postnatal enteric progenitor cells retain some characteristics of embryonic ENCDCs, but are more restricted in their self-renewal potential and their lineage commitment (Kruger *et al.*, 2002). Nevertheless, this fact makes the ENS attractive to be used as stem cell source. Respective cells are suitable to mimic disease origin and progression *in vitro* or for disease treatment in future (Burns *et al.*, 2016; Metzger *et al.*, 2009; Obermayr and Seitz, 2018; Rollo *et al.*, 2016).

1.1.3 *Signalling pathways*

Generation of a functional ENS depends on complex interactions between different cell types and cell functions, which need to be tightly controlled. Responsible signalling pathways are highly coordinated in temporal and spatial manner. Key functions that need to be orchestrated during ENS development are **cell migration, proliferation**, **survival** and **differentiation:** While vagal derived ENCDCs migrate caudally and thereby colonize the entire length of the bowel, progenitor cells start to differentiate into neurons and glial cells**.** To prevent depletion of the progenitor pool and to ensure adequate colonization of the most distal gut, part of the ENCDCs need to maintain their proliferative capacity (survival) and at the same time they have to increase in number (proliferation) (Heanue and Pachnis, 2007; Obermayr *et al.*, 2013; Sasselli *et al.*, 2012).

Except for intrinsic cell signals, secreted factors of the surrounding tissue as guiding molecules and differentiation factors, contribute to the regulation of cell functions impaired in HSCR (Heanue and Pachnis, 2007; Rao and Gershon, 2018). Central mediators involved can be grouped into three main signalling pathways, as outlined in the following (Lake and Heuckeroth, 2013).

Ret signalling pathway

The tyrosine kinase receptor Rearranged during transfection (Ret) plays a crucial role in ENS development. As a transmembrane receptor of the cadherin superfamily it binds ligands of the neurotrophic factor family, e.g. Glial-derived neurotrophic factor (Gdnf), Neurturin (Nrtn), Artemin (Artn), Persephin (Pspn). Upon ligand binding, Ret assembly with its co-receptor Gdnf family receptor alpha 1 (Gfrα1) of the glycosylphosphatidylinositol (GPI)-anchored co-receptor family is initiated. Ret signalling

plays multiple roles in ENS development (Taraviras *et al.*, 1999). During early phases of ENS formation (mouse: E9.5), Ret expression is involved in the migratory behaviour of pre-enteric NCCs. The surrounding gut mesoderm secretes Gdnf and its specific co-receptor Gfrα1, that serve as chemoattractants for Ret^+ cells. Receptor expression is maintained during gut colonization, where Ret signalling regulates proliferation and cell survival of progenitor cells. At stage E10.5 when neuronal lineage commitment is initiated, Ret is expressed in prospective enteric neurons (Hao and Young, 2009; Sasselli *et al.*, 2012). Even postnatally, a subset of enteric neurons maintains respective gene expression although detailed understanding about its role within these cells is missing so far (Rao and Gershon, 2018).

Ednrb signalling pathway

Another key mediator in ENS development is Endothelin receptor type B (Ednrb). It encodes a G-coupled receptor specific for small peptides as Endothelin 1, 2 and 3 (Edn1, Edn2, Edn3) (Bondurand *et al.*, 2018). In mouse, Ednrb starts to be expressed in ENCDCs at stage E10.5. Edn3 secretion by the surrounding gut mesenchyme is crucial for targeted migration of $Ednrb^+$ cells and for colonization of the mid- and hindgut. Highest Edn3 levels are detectable in the caecum (Heanue and Pachnis, 2007; Sasselli *et al.*, 2012). In contrast to Ret^+ cells, $Ednrb^+$ cells remain in an uncommitted state even during later phases of ENS development. Consequently, the pool of progenitor cells is maintained. Despite this antagonistic effect to the Gdnf-Ret signalling pathway, Edn3 can have a synergistic effect with regard to increased proliferation levels in Ret^+ ENCDCs (Sasselli *et al.*, 2012).

Further mediators and cell type markers

Factors modulating the activity of the Ret and Ednrb signalling pathways represent the third group of molecular mediators crucial for ENS development.

The transcriptional regulator Sex determining region Y (SRY) Box 10 (Sox10) is another major player in ENS development. Its expression in NCC-derived progenitor cells is detectable subsequently to NCC delamination in mouse at stage E8.5. Sox10 is crucial for the maintenance of multipotency and cell survival through inhibition of apoptosis in the respective cell population. While Sox10 expression is downregulated in future neurons, prospective glial cells maintain the expression (Sasselli *et al.*, 2012). Expression levels of this transcriptional regulator are associated to gliogenesis in general and neuronal subtype generation in the small intestine (Rao and Gershon, 2018). Transcriptional regulation of Ednrb and Ret is mediated through binding of Sox10 to respective enhancer binding sites (Heanue and Pachnis, 2007).

Another transcriptional regulator involved in ENS development is Paired-like homeobox 2b (Phox2b). Comparable to Ret and Sox10, Phox2b is as well expressed in early bipotent pre-enteric NCCs (Lake and Heuckeroth, 2013). Phox2b itself is involved in the regulation of Ret and Sox10 expression levels and therefore contributes to progenitor cell survival and multipotency maintenance during early phases of ENS development (Goldstein *et al.*, 2013; Rao and Gershon, 2018). At later stages, Phox2b is involved in cell differentiation, as its expression is maintained in all enteric neurons and some enteric glial cells (Sasselli *et al.*, 2012).

The Low-affinity neutrophin receptor p75 (p75) is another important mediator for ENS formation. As a member of the Tumour necrosis factor (TNF)-superfamily, it binds neutrophins as Nerve growth factor (Ngf), Brain-derived growth factor (Bdnf), Neutrophin 3 and 4 (Nt3, Nt4) (Heanue and Pachnis, 2007; Underwood and Coulson, 2008). Comparable to Sox10, it is as well expressed in early pre-enteric NCCs after delamination from the neural tube. Moreover, p75 expression is as well decreased in cells committed to the neuronal lineage, while it is maintained in glial cells (Heanue and Pachnis, 2007).

Co-expression of Phox2b, Sox10, Ret and p75 is detectable for cells at the caudal end of the migrating wavefront and keeps these cells in an uncommitted state. More cranially to this wavefront, cells are at different neuronal maturation states. In general, neuronal cell fate commitment in mouse is initiated at stage E10.5 and is associated to increased expression levels of pan-neuronal markers, such as Achaete-Scute Family BHLH (Basic Helix Loop Helix) Transcription Factor 1 (Mash1), Ubiquitin C-Terminal Hydrolase L1 (Pgp9.5), Beta III-Tubulin (Tubb3) and Hu-Antigen C/D (HuC/D). These immature neuronal cells remain mitotically active and project their outgrowing neurites to the wavefront (Goldstein *et al.*, 2013; Sasselli *et al.*, 2012). Compared to immature neurons which transiently express combinations of different neurotransmitters, distinct neuronal subtypes exit the cell cycle and are then immunoreactive to specific markers only. Subtype specific markers, indicating the presence of serotonergic and nitridergic neurons within the enteric ganglia, are first detectable between E11.5-E12.5. ACh, Vip, Neuropeptide Y (Npy) and Substance P (Tac1) positive neurons are generated subsequently between E14-E18. In general, neuronal subtype generation occurs heterogenous- and asynchronously and continues postembryonically (Hao and Young, 2009; Sasselli *et al.*, 2012). Cell commitment to the glial lineage is connected with increased expression levels of the glial specific markers S100 Calcium binding Protein b (S100b), Glial fibrillary acidic protein (Gfap) and Fatty-acid-binding protein (Fabp) (Goldstein *et al.*, 2013). Glial cell differentiation is initiated at E11.5, but mature $Gfap^{+}$ cells are not detectable before E16.5 (Rothman *et al.*, 1986).

Disturbances in the presented molecular networks controlling ENS development can cause neurocristopathies, as for example Hirschsprung's disease (HSCR).

1.2 Hirschsprung's disease

1.2.1 *Clinical characteristics*

The neurocristopathy Hirschsprung's disease (HSCR) has been first described by the Dutch physician Harald Hirschsprung in 1888 (Hirschsprung, 1888). HSCR represents the main cause for neonatal intestinal obstruction and is also referred to as congenital aganglionic megacolon. It is a developmental disorder of the ENS characterized by the complete absence of enteric ganglia in distinct segments of the distal bowel (Amiel *et al.*, 2008). Depending on the length of the affected segment, the severity ranges between the most prevalent short segment form (S-HSCR: 80%), which is limited to the rectum and the sigmoid colon, a long segment form (L-HSCR: up to 20%), which extends further to the more proximal colon, and a total colonic aganglionosis (TCA: up to 5 %) (Heanue and Pachnis, 2007; Parisi, 2015). In the majority of infants, the disease manifests within the first days after birth, as the ability to pass stool (meconium) is perturbed. Only in a minority of patients the diagnosis is made in later childhood (Parisi, 2015). On clinical level, the aganglionosis causes a tonic contraction of the affected region while the more proximal part extends massively. Consequently, a megacolon is formed, and patients suffer from enterocolitis, constipation, abdominal pain and emesis (Amiel *et al.*, 2008; Heanue and Pachnis, 2007). On histological level, HSCR is diagnosed by the lack of enteric ganglia in biopsies of the gut wall. Surgical resection of the aganglionic segment via "pull-through" technique and the reconnection of the non-affected bowel to the anus represents the only treatment option for all patients (Parisi, 2015). Functional outcomes vary among patient groups (depending on the severity degree) and among the pull-through procedure that has been applied. Even though detailed understanding about postoperative impairments is still missing, a significant number of patients continues to suffer from GI impairments as enterocolitis (25-45%) and constipation (3-22%) (Calkins, 2018; Faure, 2017).

1.2.2 *Genetics and research*

HSCR is defined as a rare genetic disease with an incidence of 1:5000 live births and shows a male sex predominance with a ratio of 4:1. In about 80% of patients, HSCR occurs as an isolated form. Only in a minority of patients, HSCR is associated to other syndromes as trisomy 21, Mowat-Wilson and Multiple endocrine neoplasia type 2A (MEN2A) and Waardenburg-Shah being the most frequent ones. Additionally, HSCR can be associated to other congenital abnormalities as heart or genitourinary malformations. These abnormalities might be caused by defects in other NC derivatives and occur most frequently in familial cases of HSCR. The mode of inheritance for most of the HSCR risk genes is autosomal dominant in familial cases. However, in case of the syndromic form the inheritance pattern is highly variable ranging from autosomal dominant or recessive to X-linked recessive (Amiel *et al.*, 2008; Goldstein *et al.*, 2013; Parisi, 2015).

So far, about 25 risk genes have been replicated to be causative for HSCR, but many more genetic loci were shown to be involved in shaping the phenotypic variability. More recently, gene-environment interactions were as well demonstrated to have an impact on ENS development and hence to contribute to the aetiology of enteric neuropathies. Therefore, HSCR is classified as multifactorial trait. Moreover, it shows reduced penetrance and variable expressivity. However, mutations in the so far identified HSCR

risk loci only account for about 30% of all patients (Bahrami *et al.*, 2017; Heuckeroth and Schäfer, 2016; Luzón-Toro *et al.*, 2015b; Tilghman *et al.*, 2019).

Risk genes can be mainly assigned to signalling cascades that are of importance in early ENS development. Consequently, the major susceptibility locus so far described for HSCR is *RET*. Rare and common variants in the coding, and as well in the non-coding region of the gene, have already been described in the context of HSCR. Rare coding variants are more frequently found in patients affected by the long-segment form while common variants often occur in short-segment phenotypes. In general, mutations in the *RET* gene account for about 50% of the familial and 15-35% of the sporadic cases. On functional level, the majority of genetic variants result in a loss-of-function of this tyrosine transmembrane kinase receptor (Heanue and Pachnis, 2007; Lake and Heuckeroth, 2013; So *et al.*, 2011). Gain-of-function mutations within this gene are implicated in the thyroid cancer syndrome MEN2A (Parisi and Kapur, 2000). In about 5 % of HSCR cases, mutations in other genes than ***RET*** have been identified. These are: ***GDNF***, ***NRTN***, ***PSPN***, ***GFRα1***, ***EDNRB***, ***EDN3***, *(ENDOTHELIN CONVERTING ENZYME 1)* ***ECE1***, ***SOX10***, ***PHOX2B***, *ZINC FINGER E-BOX BINDING HOMOEOBOX (**ZFHX1B**), L1 CELL ADHESION MOLECULE (**L1CAM**), KIF1 BINDING PROTEIN (**KIAA1279**), TRANSCRIPTION FACTOR 4 (**TCF4**), PROKINETICIN 1 (**PROK1**), PROKINETICIN RECEPTOR 1* and *2 (**PROKR1**, **PROKR2**), NEUTROPHIN 3 (**NTF3**), NEUROTROPHIC RECEPTOR TYROSINE KINASE 3 (**NTRK3**), NEUREGULIN 1* and *3 (**NRG1**, **NRG3**), CLASS 3 SEMAPHORINs (**SEMA3A**, **SEMA3C*** and ***SEMA3D**)* and *DNA METHYLTRANSFERASE 3 BETA (**DNMT3B**)* (Luzón-Toro *et al.*, 2015b).

In total, approximately 500 rare coding variants residing in the major HSCR susceptibility loci have been already described. Except for these rare coding variants, a number of common non-coding single nucleotide polymorphisms (SNPs) were associated with increased HSCR risks. It was shown that the risk alleles synergistically affect key ENS signalling pathways. Among them, the transcription-enhancer variant rs2435357 in the first intron of *RET* was described to significantly increase the HSCR risk (Chatterjee *et al.*, 2016; Emison *et al.*, 2010; Emison *et al.*, 2005; Tilghman *et al.*, 2019). Another layer of genetic complexity is added by the fact that not only syndromic cases (e.g. HSCR patients with Down Syndrome) but as well non-syndromic HSCR subjects frequently present with structural genetic variations. Therefore, the presence of rare copy number variations (CNVs) might very likely as well contribute to the phenotype (Tang *et al.*, 2012; Tilghman *et al.*, 2019).

The relevance of the previously listed HSCR risk genes were so far mainly assessed in animal models, as described in the following: A multitude of different *Ret* mutant mouse strains was generated to study the role of this gene for HSCR in more detail. Targeted gene knockout (KO) of *Ret* in mice causes total intestinal aganglionosis accompanied by kidney defects. *Ret*$^{-/-}$ mice die at birth. Less severe phenotypes ranging from hypoaganglionosis and colonic aganglionosis are caused either by mono-isoformic mutant alleles or mutated phosphorylation sites (Bondurand and Southard-Smith, 2016). Investigations in the zebrafish confirmed *ret* as major HSCR susceptibility locus, since morpholino mediated knockdown in embryos lead to the absence of enteric neurons within the distal gut (Heanue and Pachnis, 2007).

Targeted gene KOs of the Ret ligand *Gdnf* or its co-receptor *Gfrα1* in mice cause almost identical phenotypes as observable for *Ret* mutant strains. This is in line with different missense mutations in the *GDNF* gene that were identified in HSCR patients (Angrist *et al.*, 1996). Less severe phenotypes are

observable when other Ret signalling components are targeted by gene-specific deletions. Consequently, *Nrtn* deficient mice show a reduced enteric innervation and impairments in GI motility but are viable and fertile (Bondurand and Southard-Smith, 2016; Heuckeroth *et al.*, 1999). Doray *et al.* identified a heterozygous mutation in the *NRTN* gene, which was not disease-causing on its own but segregated with a *RET* mutation. Therefore, they concluded an oligogenic inheritance pattern, as mutations in further *RET*-associated signalling factors might act in concert with an additional *RET* variant (Doray *et al.*, 1998).

Corresponding to its role in early ENS development, *EDNRB* is another validated HSCR risk gene as the deletion of the receptor or point mutations within its ligand *Edn3* cause an aganglionosis in the distal colon in mice. Moreover, such mutant strains display pigmentation defects due to dysfunctions in NCC-derived melanocytes. In this manner, the syndromic Waardenburg-HSCR phenotype in humans is reflected. Mutations within *EDNRB* or its associated signalling factors *EDN3* and *ECE1* account for up to 5% of all HSCR cases (Bondurand and Southard-Smith, 2016; Heanue and Pachnis, 2007).

A HSCR resembling phenotype in mice is also observable in the case of targeted mutations and gene KOs of *Sox10*, *Phox2b* or *Zfhx1b*. For example, specific *Phox2b* and *Sox10* homozygous mutants display a complete intestinal aganglionosis (Bondurand and Southard-Smith, 2016; Heanue and Pachnis, 2007).The relevance of these two genes was also confirmed in an additional *in vivo* system as respective zebrafish mutants lack the ENS as well (Shepherd and Eisen, 2011). In humans, genetic variations within *SOX10*, *PHOX2B* and *ZFHX1B* were almost exclusively identified in syndromic HSCR subjects. Corresponding to this, haploinsufficiency of *SOX10* in HSCR patients is associated with Waardenburg syndrome, as it was initially reported by Pingault *et al.* in 1998 (Amiel *et al.*, 2008; Heanue and Pachnis, 2007; Pingault *et al.*, 1998).

Except for single gene contributions in HSCR, animal models as well reflect the multifactorial disease background. With regard to this, e.g. epistatic interactions between the major signalling pathways *Ret* and *Ednrb* or *Ednrb* and *Sox10* were already described to affect the phenotypic HSCR severity in the corresponding *in vivo* model (Cantrell *et al.*, 2004; McCallion *et al.*, 2003; Touré *et al.*, 2019).

Although animal models can mimic certain human HSCR traits, there are limitations on geno- and phenotype level. Most mouse strains display a HSCR phenotype only in the homozygous mutation state, while patients usually harbour only heterozygous variants in the respective genes. Next, many animal models display various other phenotypic alterations besides the ENS-specific impairment. These additional phenotypes are not always present in patients. Moreover, a sex predominance is mostly not displayed by any *in vivo* model (Bondurand and Southard-Smith, 2016; Burns *et al.*, 2016; Heanue and Pachnis, 2007; McCallion *et al.*, 2003). To circumvent these limitations, more recent studies made use either of patient-specific primary ENS progenitor cells or induced pluripotent stem cells (iPSCs) for general functional *in vitro* cell characterization and to analyse HSCR risk variants (Fattahi *et al.*, 2016; Lai *et al.*, 2017; Rollo *et al.*, 2016). The use of suitable *in vitro* approaches is very promising with regard to the examination of underlying mechanisms and the development of cell-based treatment strategies for enteric neuropathies in future. However, more detailed investigations and further validation procedures are necessary to overcome current challenges e.g. concerning cell isolation, propagation and characterization (Burns *et al.*, 2016; Obermayr and Seitz, 2018; Schmitteckert *et al.*, 2019).

Besides the functional characterization of HSCR risk loci and variants, dissection of the genetic heterogeneity of this enteric neuropathy is another aim targeted in current research projects. Recent studies focused on the identification of additional disease-causing variants or novel HSCR susceptibility loci by next-generation sequencing analyses. These investigations were combined with various bioinformatic filtering strategies suitable for gene and variant prioritization (Gui *et al.*, 2017; Wu *et al.*, 2018; Zhang *et al.*, 2017). To unravel the genetic architecture in familial HSCR cases, DNA sequencing analyses were coupled with segregation and linkage analysis (Luzón-Toro *et al.*, 2015b). However, only in a minority of studies, novel genetic findings were confirmed for their HSCR relevance on experimental level using e.g. the zebrafish as *in vivo* model system (Gui *et al.*, 2017; Tilghman *et al.*, 2019).

1.3 Background of the project and preliminary work

As outlined previously, HSCR is classified by a vast heterogeneity on genotype and phenotype level, which can only partially be addressed in *in vivo* model systems. Unknown genetic causes in the majority of HSCR individuals, together with limited therapy options and lifelong persisting postoperative GI impairments in many patients request for a better understanding about the molecular mechanisms underlying this neurodevelopmental disorder. Although dissection of the individual's genetic background might help to get further insights in the genetic heterogeneity of HSCR, validation of genetic findings on functional level remains a major challenge which is still often neglected in current studies.

This PhD thesis is based on preliminary work of two previous Master Thesis projects, which aimed to establish a complementary study pipeline, based on two major parts (A and B), in an initial pilot study setup.

(A) Identification and selection of novel HSCR candidate genes was performed, as described in a previous project (Master Thesis Volz, 2014).

In brief, whole exome sequencing (WES) was performed on the DNA of two sporadic L-HSCR cases and their non-affected family members. Unaffected relatives served as controls for bioinformatics filtering of sequencing data. After variant calling, synonymous as well as variants with a minor allele frequency (MAF) >1% in the 1000 Genomes Project or in Exome Aggregation Consortium (ExAC) database were removed. Next, variants were filtered for the disease model by prioritizing *de novo*, homozygous, hemizygous and compound heterozygous variants for follow up analyses; 19 and 20 novel HSCR candidate genes were identified for family I and II, respectively. None of these genes was already reported in the context of HSCR or ENS development. To nail down the lists of putative candidates, a network analysis using the Ingenuity pathway analysis (IPA) software was performed. Thereby, a minority of candidates was identified to display multiple indirect interactions to already known HSCR risk genes. Based on these findings, two candidate genes per family were selected for further validation: ***ATP7A*** (*ATPase COPPER TRANSPORTING ALPHA*) and ***SREBF1*** (*STEROL REGULATORY ELEMENT BINDING TRANSCRIPTION FACTOR 1*) represented most promising candidate genes for family I, while ***ABCD1*** (*ABC SUBFAMILY MEMBER 1*) and ***PIAS2*** (*PROTEIN INHIBITOR OF ACTIVATED STAT 2*) were the most promising genes for family II. *ATP7A, SREBF1, ABCD1* and *PIAS2* were further investigated for their potential contribution to ENS development and HSCR-causing

pathomechanisms by a web-based literature search. In this manner, further evidence for a putative relevance of these genes in HSCR accumulated. Next, gene-specific variants in all four candidate genes detected by WES were validated by Sanger sequencing. Additionally, PCR based analyses in murine tissues containing pre-migratory vagal and sacral NCCs at stages E8.5, E9.5 and E10.5 confirmed respective candidate gene expression in prospective ENS tissue (Master Thesis Volz, 2014).

(B) Detailed functional characterization of selected novel HSCR candidate genes was initiated in a previous project (Master Thesis Mederer, 2015).

For the experimental validation of *ATP7A, SREBF1, ABCD1* and *PIAS2* as promising candidate genes for HSCR, protein expression analyses in murine tissues of different embryonic developmental stages were carried out. Srebf1 and Pias2 were prioritized in the respective analyses and were found to be expressed within the developing GI system. For more detailed functional *in vitro* analyses targeting the neuronal specific relevance of the candidate genes, the generation of a suitable human neuronal cell culture model system was envisioned. Therefore, the CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)/Cas9 technology was selected for genome editing in the neuroblastoma cell line SHSY5Y. It was aimed to generate gene-specific *KO* cell clones for all candidate genes and the major HSCR susceptibility locus *RET* which should be characterized on various levels in comparative *in vitro* analyses. Genome edited clones for *PIAS2* and *RET* were already generated but not validated for further usage (Master Thesis Mederer, 2015).

1.4 Aim of the project

Taking previous data into account, this project focused on the establishment of a complementary research approach for HSCR based on genetic, bioinformatics, molecular and functional data.

In a first step, it was aimed to confirm the initial candidate selection presented in 1.3 (part A). Therefore, WES data was re-evaluated on the basis of current literature and database states. Additionally, the patient's genetics were dissected in more detail to identify rare coding variants in ENS related genes, common non-coding HSCR risk variants and CNVs. Transcriptome wide mRNA profiling of different embryonic murine ENS-relevant tissues was performed for experimental validation of the initial candidate gene selection. Additionally, further rare variant carriers for the candidate genes were identified to gain more genetic based evidence for their selection and strengthen their putative relevance for HSCR.

In a second step, it was aimed to assess the candidate genes for their relevance in the HSCR aetiology in detail. To this end, candidate expression analyses on protein level, as outlined in 1.3 (part B), were extended within this project, as further tissues, developmental stages and cell types were analysed. The previously established genome editing strategy for the neuroblastoma cell line was modified and applied to all candidate genes. Moreover, a neuronal-like cell culture model system suitable for comparative analyses of the generated *KO* cell clones was established and applied. Various functional *in vitro* assays were implemented in the study pipeline and used for detailed characterization of the candidate genes.

To allow a final risk evaluation and classification of the four candidate genes for HSCR, generated data was summarized in a binary scoring system which was established in another study (Bachelor Thesis Dawid, 2019).

2 MATERIAL AND METHODS

2.1 Material

2.1.1 *Bacterial strains, cell lines and mouse line*

Bacterial strains

Chemocompetent *DH5α E.Coli*

[*supE*44 Δ*lac*U169 (*ϕ*80 *lacZ*ΔM15) *hsdR*17 *recA*1 *endA*1 *gyrA96 thi-1 relA1*]

Chemocompetent XL10 Gold *E.Coli* (Agilent)

[Tet^r *Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte*

[F´*proAB lacIqZΔM15 Tn10* (Tetr) Amy Cam^r]

Chemocompetent NovaBlue Singles Competent Cells (Merck Millipore)

[*endA1 hsdR17* (r_{K12^-} m_{K12^+}) *supE44 thi-1 recA1 gyrA96 relA1 lac* F'[*proA*$^+$*B*$^+$ *lacI*^{q}*ZΔM15*::Tn*10*] (Tet^R)]

Cell lines

Human embryonic kidney cells (HEK293TN)	Biocompare
Human epithelial neuroblastoma bone marrow cells (SHSY5Y)	DSMZ

Mouse line

CD1 wildtype (WT) mice were housed in specific pathogen free (SPF) conditions at the Interfacultary Biomedical Faculty (IBF) (University of Heidelberg). Mice were fed *ad libitum*. The day of vaginal plug was defined as E0.5. CD1 WT mice were sacrificed by CO_2 asphyxiation.

2.1.2 *Reagents for bacterial culture*

Ampicillin	Roth
Bacto-Agar	BD
Bacto-Tryptone	BD
Glucose	Merck
Isopropyl β-D-1-thiogalactopyranoside (IPTG)/ X-Gal	usb
Kanamycin	Roth
Magnesium chloride ($MgCl_2$)	Merck
Magnesium sulfate ($MgSO_4$)	Merck
Potassium chloride (KCl)	Merck
Sodium chloride (NaCl)	Sigma-Aldrich
Yeast extract	BD

2.1.3 *Media for bacterial cultures*

LB-Agar (ampicillin/kanamycin)

LB medium + 1.5% Bacto-Agar + 50µg/mL ampicillin or 30µg/ml kanamycin

LB-medium (ampicillin/kanamycin)

1% Bacto-Tryptone + 0.5% Yeast extract + 1% sodium chloride + 50µg/mL ampicillin or 30µg/ml kanamycin

SOB-medium

10mM sodium chloride + 10mM magnesium chloride + 0.5mM potassium chloride + 10mM magnesium sulfate + 2% Bacto-Tryptone + 0.5% Yeast extract

SOC-medium

1L SOB-medium + 20mL glucose (20%)

2.1.4 *Reagents for cell culture*

Acetic acid	Merck
Brain-derived neurotrophic factor (BDNF)	Peprotech
Collagen I (Col I)	Thermo Fisher Scientific
Collagen IV (Col IV)	Sigma-Aldrich
Distilled water	Thermo Fisher Scientific
Dulbecco's modified eagle medium (DMEM) (high glucose)	Thermo Fisher Scientific
Dulbecco's phosphate buffered Saline (PBS)	Thermo Fisher Scientific
Fetal bovine serum (FBS)	Thermo Fisher Scientific
Laminin	Sigma Aldrich
Matrigel	Corning
Neurobasal medium, without L-Glutamine	Thermo Fisher Scientific
Neuregulin1 (NRG1)	Peprotech
Neurotrophic growth factor (NGF)	Peprotech
Penicillin/Streptomycin	Thermo Fisher Scientific
Retinoic acid (RA)	Sigma-Aldrich
StemPro Accutase	Thermo Fisher Scientific
Vitamin D3 (Cholecalciferol)	Sigma-Aldrich
0.05% Trypsin-EDTA	Thermo Fisher Scientific

2.1.5 *Media and buffers for cell culture*

MACS buffer

1X PBS +1% FBS + 2.5mM EDTA

HEK293TN culture medium (HEK293TN CM)

DMEM (high glucose) + 10% FBS

SHSY5Y culture medium (SHSY5Y CM)

DMEM (high glucose) + 15% FBS + 1% Pen/Strep

SHSY5Y differentiation medium I (DM I)

DMEM (high glucose) + 15% FBS + 1% Pen/Strep + 10µM RA

SH-SY5Y differentiation medium II (DM II)

DMEM (high glucose) + 1% Pen/Strep + 50ng/µL BDNF

SHSY5Y differentiation medium III (DM III)

DMEM (high glucose) + 1% Pen/Strep + 15% FBS + 10µm RA + 50ng/µL BDNF + 10ng/µL NRG1 + 10ng/µL NGF + 24nM Vitamin D3

2.1.6 *Vector constructs*

pSpCas9(BB)-2A-GFP (PX458) (#48138)	Addgene
pcDNA3.1-2xFLAG-SREBP-1c (#26802)	Addgene
pcDNA3-CASP3-MYC (#11813)	Addgene
pcDNA4-ATP7A-MYC	kindly provided by Bart van de Sluis University of Groningen, The Netherlands
pCMV_Flag_PIASxalpha (#15209)	Addgene
pEGFP-N1-ABCD1	kindly provided by Dr. Imanaka, University of Toyama, Japan

2.1.7 *Reagents for molecular analyses*

Acetic acid	Merck
LE Agarose	Biozym
Ammonium persulfate	GR Healthcare
AquaPolymount	Polysciences
Bovine Serum Albumin (BSA) – Fraction V, pH 7.0	Serva Electrophoresis
Bromphenol blue	Sigma-Aldrich
Chloroform	Sigma-Aldrich
Citric acid	Merck
Diethly pyrocarbonate (DEPC)	Sigma-Aldrich
Dimethylsulfoxide (DMSO)	Sigma-Aldrich

DirectPCR Lysis-Reagent Cell	Peqlab
Ethanol (EtOH)	Sigma-Aldrich
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich
GeneRuler DNA Ladder (100bp /1 kb)	Thermo Fisher Scientific
Glycerol	Sigma-Aldrich
Glycine	Sigma-Aldrich
Hoechst 33342 trihydrochloride, trihydrate	Thermo Fisher Scientific
Hydrochloric acid (37%)	Merck
Igepal (NP-40)	Chem-Cruz
Immobilon FL-Polyvinylidene fluoride (PVDF) Membrane	Merck
Isopropanol	Sigma-Aldrich
Lipofectamine 2000	Thermo Fisher Scientific
Methanol (MeOH)	Sigma-Aldrich
Midori Green DNA stain	Nippon Genetics
N,N,N′,N′-Tetramethylethylenediamine (TEMED)	Sigma-Aldrich
Normal goat serum (NGS)	Thermo Fisher Scientific
Odyssey blocking reagent (PBS)	LI-COR Biosciences
Opti Minimal Essential Medium (OptiMEM)	Thermo Fisher Scientific
Paraformaldehyde (PFA)	Sigma-Aldrich
Proteinase and phosphatase inhibitor cocktail	Thermo Fisher Scientific
Rotiphorese Gel 30 (37,5:1) Acrylamid-Bisacrylamid solution	Carl Roth
Sodium chloride	Sigma-Aldrich
Sodium citrate	AppliChem
Sodium deoxycholate	Sigma-Aldrich
Sodium dodecyl sulphate (SDS)	Sigma-Aldrich
Sucrose	AppliChem
Tissue freezing medium	Jung
TRIS (Tris-(hydroxymethyl)-aminomethan)	Carl Roth
Triton X-100	Sigma-Aldrich
Trizol	Thermo Fisher Scientific
Tween-20	Carl Roth
Vectashield antifade mounting medium for fluorescence	Vector Laboratories

2.1.9 *Buffers for molecular analyses*

All buffers were prepared with deionized water.

50X TRIS-Acetate-EDTA (TAE) buffer (DNA gelelectrophoresis buffer)

2M TRIS + 1M acetic acid + 50mM EDTA

5X Laemmli Running buffer (Western blot)

125mM TRIS + 960mM Glycine + 0.5% SDS

10X PBS

13.7M sodium chloride + 270mM potassium chloride + 800mM disodium hydrogenphosphate + 200mM potassium dihydrogen phosphate (pH 7.4)

Citrate buffer

1.8mM Citric acid + 8.2mM Sodium Citrate + 0.05% Tween-20 (pH 6)

Radio Immuno Precipitation Assay (RIPA) buffer

150mM sodium chloride + 1% NP-40 + 0.5% sodium deoxycholate + 0.1% SDS + 50mM TRIS (pH 8)

Stacking Gel buffer (Western blot)

1.5M TRIS (pH 8.8)

Separating Gel buffer (Western blot)

1M TRIS (pH 6.5)

10X TBS

200mM TRIS + 1.37M sodium chloride

TBST

1X TBS + 0.1% Tween-20

10X Transfer buffer (Western blot)

250mM TRIS + 1.92M Glycin

1X Transfer bufffer

1X Transfer buffer + 10% MeOH

2.1.10 *Kits*

Bioanalyzer RNA analysis Kit	Agilent
BrdU (Bromodeoxyuridine) Incorporation Kit	Roche
CombiMag	OZ Biosciences
Direct PCRLysis Reagent	Peqlab
FastDigest Restriction Enzymes *BamHI*, *BbsI*, *SrfI*, *XbaI*	Thermo Fisher Scientific
GeneJET Plasmid MiniPrep Kit	Thermo Fisher Scientific
HotStarTaq DNA Polymerase Kit	Qiagen
In Situ Cell Death Detection Kit, Fluorescein	Roche
Pierce BCA Protein Assay Kit	Thermo Fisher Scientific
pSTBlue-1 AccepTor Vector Kit	Merck Millipore
Q5 High fidelity DNA Polymerase Kit	New England Biolabs
QIAquick Gel Extraction Kit	Qiagen
QIAquick PCR Purification Kit	Qiagen
Qubit RNA BR Assay Kit	Thermo Fisher Scientific
Quick DNA Microprep Kit	Zymo Research
QuickChange Lightning Site-Directed Mutagenesis Kit	Agilent
RNAqueous-Micro Total RNA Isolation Kit	Thermo Fisher Scientific
SensiFAST SYBR Lo-ROX Kit	Bioline
Superscript III First-Strand Synthesis System for reverse transcription PCR	Thermo Fisher Scientific
T4 DNA Ligase	Thermo Fisher Scientific
T4 PNK Kit	Thermo Fisher Scientific
ZymoPURE II plasmid prep Kit	ZymoResearch

2.1.11 *Antibodies*

Table 1: Primary antibodies
Primary antibody information and dilutions used for different purposes are listed below. Details about the permeabilization detergent and demasking used for immunofluorescence (IF) analysis of cryosections and for immunohistochemical (IHC) staining of formalin-fixed paraffin embedded (FFPE) sections are annotated. – not applicable

Antibody	Reactivity	Supplier	Annotations (dilution, permeabilization, demasking)			
			IF of		IHC of	
			cells	cryosections	FFPE sections	Western blot
Rabbit anti-ABCD1	Mouse, rat, human	Abcam, ab197013	1:100	1:100 (0.1% Tween)	1:100	1:5000
Rabbit anti-ATP7A	Mouse, rat, rabbit, human	Abcam, ab125137	1:75	1:75 (0.1% Triton)	-	1:250
Rabbit anti-ATP7A	Mouse, rat, human	Elabscience, E-AB-16268	-	-	1:75	-
Mouse anti-ECad	Mouse, rat, horse, human	Abcam, ab76055	-	1:100 (demasking; 0.1% Triton)	-	-
Mouse anti-FLAG		Sigma Aldrich, F3165	1:100	-	-	-
Mouse anti-GAPDH	Mouse, rat, human, chicken, xenopus, dog, fish, pig,	Abcam, ab8245	-	-	-	1:2000
Mouse anti-GAPDH	Mouse, rat, human, pig, rabbit	NovusBio, NB300-328SS	-	-	-	1:10000
Rabbit anti-GAP43	Mouse, rat, human	Abcam, ab75810	1:100	-	-	-
Mouse anti-GFP	-	Sigma Aldrich, G6539	1:100	-	-	-
Mouse anti-HuC/D-Biotin	Mouse, human	Molecular Probe, A-21272	-	1:100 (0.1% Triton)	-	-
Rabbit anti-Ki67	Mouse, rat, human, marmoset	Abcam, ab16667	1:100	-	-	-
Rabbit anti-MAP2	Mouse, rat, human, monkey	Cell signalling, 4542	1:35	-	-	-

Mouse anti-MYC	-	Cell signalling, 2276	1:100	-	-	-
Mouse anti-PGP9.5	Mouse, rat, human, rabbit, sheep, zebrafish	Abcam, ab8189	1:100	1:100 (0.1% Triton)	-	-
Rabbit anti-PIAS2	Mouse, rat, human, chicken	Biozol, GTX115180	1:100	1:100 (0.1% Triton, demasking)	1:500	-
Goat anti-PIAS2	Human, predicted: mouse, rat, dog	Biozol, ARG64201.100	-	-	-	1:500
Rabbit anti-P75	Mouse, rat, human	Upstate, 07-476	-	-	-	-
Rabbit anti-RET	Mouse, rat, human	Abcam, ab134100	1:40	1:100 (0.1% Triton)	-	1:1000
Mouse anti-SMA	Mouse, rat, human, Bovine, chicken, rabbit, baboon	Abcam, ab18147	-	1:250 (0.1% Triton)	-	-
Mouse anti-SOX10	Mouse, rat, human	Affymetrix, 14-5923	-	1:50 (0.1% Triton)	-	-
Rabbit anti-SREBF1	Mouse, rat, human	Abcam, ab28481	1:300	-	1:1000	--
Rabbit anti-SREBF1	Mouse, rat, human, pig, hamster	Abnova, PAB12021	-	1:300 (1% Triton)	-	1:500
Rabbit anti-SYP	Mouse, rat, human	Abcam, ab32127	1:100	-	-	-
Mouse anti-TAU	Mouse, rat, human, predicted: cow	Cell signalling, 4019	1:150	-	-	-
Mouse anti-TUBB3	mammals	Promega, G7121	1:500	1:500 (0.1% Triton)	-	-

Table 2: Secondary antibodies
Secondary antibody information and dilutions are listed below. – not applicable

Antibody	Conjugate	Supplier	Annotations (dilution)	
			Immuno-fluorescence	Western blot
Goat anti-mouse IgG	Alexa Fluor 488	Thermo Fisher Scientific (A-11029)	1:250	-
Goat anti-mouse IgG	Alexa Fluor 647	Thermo Fisher Scientific (A-21236)		
Goat anti-rabbit IgG	Alexa Fluor 568	Thermo Fisher Scientific (A-11011)		
Goat anti-rabbit IgG	Alexa Fluor 647	Thermo Fisher Scientific (A-21245)		
Streptavidin	Alexa Fluor 488	Thermo Fisher Scientific (S32354)		
Donkey anti mouse	IRDyeCW680	LI-COR (926-32222)	-	1:10000
Donkey anti goat	IRDyeCW680	LI-COR (932-32214)		
Donkey anti rabbit	IRDyeCW680	LI-COR (926-32223)		
Donkey anti mouse	IRDyeCW800	LI-COR (926-32212)		

2.1.12 *Oligonucleotides*

Oligonucleotides for conventional PCR were designed with *Primer3* and *OligoAnalyzer 3.1*. All primers were designed to be annealed to the target sequence at a temperature of 60°C. Oligonucleotides used for quantitative realtime PCR (qRT PCR) were designed with the *Universal ProbeLibrary Assay Design Center* from Roche. Mutagenesis primers were designed by using the web-based *QuickChange Primer Design tool* from Agilent.

Table 3: Oligonucleotides
Applications are annotated. Mutated bases are highlighted in bold. Overhang sequences for cloning are lowercased, restriction sequences for cloning are lowercased and underscored.

Primer name	Sequence (5' > 3')	Application
ABCD1_for	ATGAAGGAGACAGCAGCACC	Endogenous expression in SHSY5Y cells
ABCD1_rev	AGGGAATCATGGGGACTCGA	
ABCD1_Ex1_for	ACTTCTCCCAGCAGACCTACTA	Alternative splicing of edited clone
ABCD1_Ex3_rev	GCTCCTCCTCCTTCTTTTCCAA	
ABCD1_MUT_for	GCGCTTCAGC**C**TGTTCCCCCGCCAG	Mutagenesis experiment
ABCD1_MUT_rev	CTGGCGGGGGAACA**G**GCTGAAGCGC	

ABCD1_qPCR_for4	TTCCTCATGAAGTATGTGTGGAG	qRT PCR
ABCD1_qPCR_rev4	CCTTCTTCACGGCCTCTG	
ATP7A_for	CCCTTGTGAGTGCCTGGTAG	Endogenous expression in SHSY5Y cells
ATP7A_rev	AGAACTCCATCCCCTACCCC	
ATP7A_Ex2_for	TTGGACCATTGAGCAGCAGA	Alternative splicing of edited clone
ATP7A_Ex4_rev	GCTGGAGGGAGAGTTTGAGG	
ATP7A_MUT_for	TTAGAGCAAAGACAAAATTTATC**T** GAATCCTCTTGACTGTCTTTCTT	Mutagenesis experiment
ATP7A_MUT_rev	AAGAAAGACAGTCAAGAGGATTC**A**G ATAAATTTTGTCTTTGCTCTAA	
ATP7a_qPCR_for3	TTTGCTGTCCTTTTTATTGTGTG	qRT PCR
ATP7a_qPCR_rev3	GTGCTTTATAAGCCTGAATGTAGAAG	
CASP3_qPCR_for	TGGAATTGATGCGTGATGTT	qRT PCR
CASP3_qPCR_rev	TCCAAAAATTATTCCTTCTTCACC	
CRISPR_Cas9_rev	GTCTGCAGAATTGGCGCAC	Sequence validation sgRNA cloning
CRISPR_ABCD1_for	TTGAGTTTGAGACCTGGCCC	Genotyping of CRISPR clones
CRISPR_ABCD1_rev	CCACTTGAGCCTGGGAAGTT	
CRISPR_ATP7A_for	TGGTAGAGTGTGAAGTGTTGGAG	Genotyping of CRISPR clones
CRISPR_ATP7A_rev	CCATGAATTGCCAACCCAGC	
CRISPR_PIAS2_Ex2_T1_for	GCTTGAGGGTGACTTGGCTT	Genotyping of CRISPR clones
CRISPR_PIAS2_Ex2_T1_rev	CGGCCAAGTCAGGTTCTACA	
CRISPR_PIAS2_Ex2_T2_for	TGTGGAGGAAATGTTGCCCT	Genotyping of CRISPR clones
CRISPR_PIAS2_Ex2_T2_rev	GGCTAGTGGGAGAGGAGGAG	
CRISPR_PIAS2_Ex2_T3_for	TGATCCTCTTGTGTTGGCCTC	Genotyping of CRISPR clones
CRISPR_PIAS2_Ex2_T3_rev	AGAAACTGGAACCCTTGTGCA	
CRISPR_PIAS2_Ex6_for	TCCCCTCACTTCCCAGTTTG	Genotyping of CRISPR clones
CRISPR_PIAS2_Ex6_rev	TAACTTTAAGAATCTCCACACTGC	
CRISPR_RET_for	AAGCTGTATGTGGACCAGGC	Genotyping of CRISPR clones
CRISPR_RET_rev	CCAGCCTCACTTAACCCCTG	
CRISPR_SREBF1_Ex5_T1_for	TAACGACCACTGTGACCTCG	Genotyping of CRISPR clones
CRISPR_SREBF1_Ex5_T1_rev	ATAGGCAGCTTCTCCGCATC	
GAP43_qPCR_for	CCATGCTGTGCTGTATGAGAA	qRT PCR
GAP43_qPCR_rev	GACAGGGAAAATTCACACTTGAG	
GAPDH_for	CGACCACTTTGTCAAGCTCA	Conventional PCR (reference)
GAPDH_rev	AGGGGTCTACATGGCAACTG	

GAPDH_qPCR_for	CAGCCTCGTCCCGTAGAC	qRT PCR
GAPDH_qPCR_rev	CGCTCCTGGAAGATGGTG	
GFAP_qPCR_for	CAAGAGGAACATCGTGGTGA	qRT PCR
GFAP_qPCR_rev	GCTCCTGCTTGGACTCCTTA	
HPRT1_qPCR_for	TGATAGATCCATTCCTATGACTGTAGA	qRT PCR
HPRT1_qPCR_rev	AAACATTCTTTCCAGTTAAAGTTGAG	
MAP2_qPCR_for	CCTGTGTTAAGCGGAAAACC	qRT PCR
MAP2_qPCR_rev	AGAGACTTTGTCCTTTGCCTGT	
MASH1_qPCR_for	CGACTTCACCAACTGGTTCTG	qRT PCR
MASH1_qPCR_rev	ATGCAGGTTGTGCGATCA	
NES_qPCR_for	GAGGTGGCCACGTACAGG	qRT PCR
NES_qPCR_rev	AAGCTGAGGGAAGTCTTGGA	
NPY_qPCR_for	CTCCCCGACAGCATAGTA	qRT PCR
NPY_qPCR_rev	GCCCCAGTCGCTTGTTAC	
P75_qPCR_for	TCATCCCTGTCTATTGCTCCA	qRT PCR
P75_qPCR_rev	TGTTCTGCTTGCAGCTGTTC	
PIAS2_for	GCAAGCATGACCTCCTGATG	Endogenous expression in SHSY5Y cells
PIAS2_rev	GAAGCAGCACAGAACCAACA	
PIAS2_Exon1_for	ATGGCGGATTTCGAAGAGTTGA	Alternative splicing of edited clone
PIAS2_Exon3_rev	CTCTAACTTGTTGAGGTGTCAAAGC	
PIAS2_Exon5_for	GCAGAGACAAGTTGCCCTCA	Alternative splicing of edited clone
PIAS2_Exon7_rev	GTGCTCTGGAATGATCAGGGT	
PIAS2_MUT_for	GCTGATGTAAGCTGCC**A**TACAAGATA TACAGACATAGAGTAATTC	Mutagenesis experiment
PIAS2_MUT_rev	GAATTACTCTATGTCTGTATATCTTG TA**T**GGCAGCTTACATCAGC	
RET_for	CCTGGGAGAAGCTCAGTGTC	Endogenous expression in SHSY5Y
RET_rev	GATGTTGGGGCACAAGAACT	
RET_Exon1_for	ATGGCGAAGGCGACGTCCG	Alternative splicing of edited clone
RET_Exon3_rev	GCAGGCTGGAAAGGAGGTGTTGA	
RET_overexpr_cloning_for	gATGGCGAAGGCGACGTC	Cloning overexpression construct
RET_overexpr_cloning_rev	actgtctagaTTAACTATCAAACGTGTCCAT TAATTTTGC	
RET_rs2435357_for	GTCAAAGCCAGTGGTGATGC	Genotyping rs2435357
RET_rs2435357_rev	CAGAGTTAATCACCAGCCTCACT	
RET_rs2506030_for	CGGCCAGTTCTACCTTCTCTG	Genotyping rs2506030
RET_rs2506030_rev	AGAGCAAAACGACCCATGAGT	

RET_rs7069590_for	CTGGATCTGAGCCTGGTCTTG	Genotyping rs7069590
RET_rs7069590_rev	GGTGAGGTTAGTGGTCAGGTG	
RET3_qPCR_for	ACCACCCACATGTCATCAAA	qRT PCR
RET3_qPCR_rev	TGGCGTACTCCACGATGA	
SDHA_qPCR_for	TGGGAACAAGAGGGCATCTG	qRT PCR
SDHA_qPCR_rev	CCACCACTGCATCAAATTCATG	
SEMA_rs11766001_for	TGGGTGTGAAAGATGATGGTGT	Genotyping rs11766001
SEMA_rs11766001_rev	TGCCTCCTTTGTAGCTTCTTGT	
SOX10_qPCR_for	GACCAGTACCCGCACCTG	qRT PCR
SOX10_qPCR_rev	CGCTTGTCACTTTCGTTCAG	
SREBF1_for	CAGTGACTTCCCTGGCCTAT	Endogenous expression in SHSY5Y
SREBF1_rev	TTCCTTGATACCAGGCCCAG	
SREBF1_Exon4_for	TAACGACCACTGTGACCTCG	Alternative splicing in edited clone
SREBF1_Exon6_rev	ATAGGCAGCTTCTCCGCATC	
SREBF1_MUT_for	ACACCAGCTCCTC**C**ACCGCAGTCTGAG	Mutagenesis experiment
SREBF1_MUT_rev	CTCAGACTGC**G**GTGGAGGAGCTGGTGT	
SREBF1_FL_overexpr_cloning_for	actgggatccATGGACGAGCCACCCTTCA	Cloning overexpression construct
SREBF1_FL_overexpr_cloning_rev	actgtctagaTAGCTGGAAGTGACAGTGGT	
SREBF1_qPCR_for	CGCTCCTCCATCAATGACA	qRT PCR
SREBF1_qPCR_rev	TGCGCAAGACAGCAGATTTA	
SYP_qPCR_for	CCAATCAGATGTAGTCTGGTCAGT	qRT PCR
SYP_qPCR_rev	AGGGGTGGAGACCTAGGGTA	
TAC1_qPCR_for	GCCTCAGCAGTTCTTTGGAT	qRT PCR
TAC1_qPCR_rev	AGCCTTTAACAGGGCCACTT	
TAU_qPCR_for	GGACTGGAAGCGATGACAA	qRT PCR
TAU_qPCR_rev	GGCTAAGGCAAGGCCTATTT	
TUBB3_qPCR_for	GCAACTACGTGGGCGACT	qRT PCR
TUBB3_qPCR_rev	ATGGCTCGAGGCACGTACT	
UCHL1_qPCR_for	AGATCAACCCCGAGATGCT	qRT PCR
UCHL1_qPCR_rev	ACCGAGCCCAGAGACTCC	

Prior to single guide RNA (sgRNA) design, *in silico* analyses for annotated isoforms using the NCBI and ENSEMBL databases of respective candidate genes were performed (tab. S4, Supplementary). Exons with a base number that was not a multiple of 3 and that were located at the most 5' region of as many annotated isoforms as possible were selected as target sites. The *CCTop CRISPR/Cas9 target online predictor* was used for sgRNA design (Stemmer *et al.*, 2015). sgRNAs were re-evaluated using the *Optimized CRISPR Design portal* and the *CHOPCHOP web tool* (Montague *et al.*, 2014) (tab. S5, Supplementary). Oligonucleotides selected for genome editing are shown below. *RET* gRNAs and *PIAS2* gRNAs specific to exon 6 had been designed previously and applied in a genome editing experiment as reported in the Master Thesis Mederer, 2015.

Table 4: sgRNAs

Gene	**Exon**	**Sequence (5'>3')** (5' overhang nucleotides necessary for cloning into the vector backbone are lowercased)
RET	Ex2	caccgTCTACGGCACGTACCGCACA
		aaacTGTGCGGTACGTGCCGTAGAc
ATP7A	Ex3	caccGATCATAGTATGGCTCAAGC
		aaacGCTTGAGCCATACTATGATC
SREBF1	Ex5	caccGAGACCTGCCGCCTTCACAG
		aaacCTGTGAAGGCGGCAGGTCTC
ABCD1	Ex2	caccgCGCTCCACACATACTTCATG
		aaacCATGAAGTATGTGTGGAGCGc
PIAS2	Ex6	caccgTAATATTCAAGGGGCGTCCA aaacTGGACGCCCCTTGAATATTAc
	Ex2_T1	caccgCGATATCCACGAACTCTTGA aaacTCAAGAGTTCGTGGATATCGc
	Ex2_T2	caccgCCGAGAATTGTATAGACGCCGAT aaacATCGGCGTCTATACAATTCTCGGc
	Ex2_T2	caccgCCACGAACTCTTGAAGGACTTTC aaacGAAAGTCCTTCAAGAGTTCGTGGc

Genomic modifications introduced by CRISPR/Cas9 editing were annotated by using the *Mutationtaster* webtool. The predicted list of the top 20 off-target sites generated by the *CCTOP CRISPR/Cas9 target online predictor* was further filtered for exonic off-target sites. In addition, putatively essential intra-/intergenic off-target sites encoding e.g. non-coding RNAs or containing enhancer sites were identified using the *UCSC Genome browser*. Oligonucleotides specific to off-target sites selected for genotyping are shown below.

Table 5: Oligonucleotides used for off-target analysis

Primer name	Sequence (5‘>3‘)	Application
ARHGAP24_for	TGACGGAAGGTTAAAGGCACA	Off-targets_CRISPR *ATP7A*
ARHGAP24_rev	TTCCCCTGCTTCACTGTTGA	
DYRK1A_for	GCTTTTGTGGGTGTAAGGGC	
DYRK1A_rev	CCCAGGAGTCAGTTTGCAGT	
EGF_for	GATGACACTTGGGAGCCTGG	
EGF_rev	GGCTGGTCTCAAACTCCTGG	
NDRG3_for	GTCAGGCCAAGTGTACCCTC	
NDRG3_rev	GGCCTCAGTTACCCCATCTG	
RP11-146N_for	TGCCACGATTCTTCCTCTGG	
RP11-146N_rev	ACTGACCTACTGCCCTGCTA	
SMOX_for	GTCCTGTCTTCCCCATGCAG	
SMOX_rev	TATATGCCCCTTGGTGGTGC	
ZMAT5_for	CAACAACGCTGACTGAAGGC	
ZMAT5_rev	CCTTCCTTCTGCAACTGGGT	
ST8SIA5_rev	TCCTCCTCCCAGCACTAACA	Off-targets_CRISPR *SREBF1*
ST8SIA5_for	TTAACAGCCAAGCCACACCA	
CAMK1G_for	TCAGTAAACACGGACCGAGC	
CAMK1G_rev	TGGCATTTCTGAGCCTCCTG	
CLIC6_2_for	TTAGGTGTTTGGGTGTCGGG	
CLIC6_2_rev	CCATTGAGTCCTCCCGCAAA	
CTDSP1_for	TGTAGAGACAGAGGGCAGCT	
CTDSP1_rev	CGCTTTCCTCTCCCTCTGTG	
EEF1A2_2_for	GTTGAGGTTCAGGTCTGGGG	
EEF1A2_rev	CCGCTCTTCTTCTCCACGTT	
MIR8073_for	ATCCTGGGCTGCCTCTATCA	
MIR8073_rev	AAGCCTGGAACCTACACAGC	
XYLB_for	GAGCCCTGTGGTTTGCTTTG	
XYLB_rev	AACCGGGTTGCAGAAGATGT	
SNORA23_for	TGTATTCCCTCTCTTCAGCAGTG	Off-targets_CRISPR *ABCD1*
SNORA23_rev	ACCCAGCAAGCCTTTTGTAAAA	

2.1.13 *Hardware*

Automated Inverted Microscope DMI4000B	Leica
Biological Safety Cabinet Safe 2020	Thermo Fisher Scientific
Centrifuges	
Centrifuge 5418	Eppendorf
Biofuge fresco	Heraeus
Cryostat-microtome CM3050S	Leica
DeNovix Spectrophotometer & Fluorometer	DeNovix
Microbiological incubator	WTB Binder
Odyssey Infrared Imager	LI-COR Biosciences
Olympus BX53 Upright microscope	Olympus
Orbital Climo-shaker	Kuhner
Orbital shaker	VWR
Polytron tissue homogenizer	Polytron
QUANTUM Gel Documentation System	Peqlab
Steri Cult CO_2 incubator	Thermo Fisher Scientific
Thermal cyclers (Mastercycler pro –vapo protect)	Eppendorf
Thermomixer	Eppendorf

2.1.14 *Software*

Affinity designer 1.7.0.	Pantone
CellSens Entry 2.1	Olympus
GraphPad Prism Vers. 5	GraphPad
Illustrator for biological sequences v1.0.3	GPS
Image J	National Institute of Health
Image Studio	LI-COR Biosciences
Ingenuity Pathway Analysis Software	Qiagen
Leica Application Suite Advanced Fluorescence 4.0.0.11706	Leica
Photoshop CS4 Vers. 11.0	Adobe
Quantum Capt Vers. 15.12	Peqlab
SnapGene Viewer Vers. 3.2.1	SnapGene

2.1.15 *Websites*

CCTop CRISPR/Cas9 target online predictor: http://crispr.cos.uni-heidelberg.de/

CHOPCHOP web tool: https://chopchop.rc.fas.harvard.edu/

ENSEMBL genome browser: https://www.ensembl.org/index.html

Expert Protein Analysis System (ExPASy): http://www.expasy.org/

Genecards: https://www.genecards.org/

Genome Aggregation Database (gnomAD): http://gnomAD.broadinstitute.org/

GTEx portal: https://gtexportal.org/home/index.html

GWAS catalogue: https://www.ebi.ac.uk/gwas/

CADD SNV lookup: https://cadd.gs.washington.edu/snv

Mutationtaster: http://www.mutationtaster.org/

National Center for Biotechnology Information (NCBI): http://www.ncbi.nlm.nih.gov/

OligoAnalyzer 3.1: http://eu.idtdna.com/calc/analyzer

Online Mendelian Inheritance in Man (OMIM): https://www.omim.org/

Optimized CRISP Design portal: http://crispr.mit.edu/

Pfam database: https://pfam.xfam.org/

Primer3web (Vers. 4.0.0): http://primer3.ut.ee/

PROVEAN: http://provean.jcvi.org/seq_submit.php

QuickChange Primer Design Program: https://www.agilent.com/store/primerDesignProgram.jsp

SFARI Gene: https://gene.sfari.org/

The Human Protein Atlas: https://www.proteinatlas.org/

UCSC Genome Browser: https://genome.ucsc.edu/

UniProt: https://www.uniprot.org/

Universal ProbeLibrary Assay Design Center: https://lifescience.roche.com/en_de/brands/universal-probe-library.html#assay-design-center)

2.2 Methods

2.2.1 *In silico analyses*

Exome sequencing of HSCR patients and variant calling

Two sporadic long-segment HSCR cases were analyzed in the context of a pilot study. Recruitment of HSCR patients was performed at the Pediatrics Surgery of the University Hospital Heidelberg by Prof. Dr. Holland-Cunz and Dr. Romero. The study has been approved by the Ethical board of the University Hospital Heidelberg (S509/2012, 08.02.2013: Genomweite Sequenzierung bei Patienten mit Morbus Hirschsprung). Parents provided written informed consent before DNA sample submission. Trio WES of patients and their non-affected family members followed by bioinformatics data filtering and candidate gene selection was carried out as reported in Volz' Master Thesis in 2014 and as briefly summarized in chapter 1.3. In the context of this project, WES data was bioinformatically re-evaluated accordingly, applying the same filtering strategy. These analyses were kindly performed by Dr. Nagarajan Paramasivam (Theoretical Bioinformatics Division, DKFZ).

CNV analysis

WES data was also used for CNV analysis. To increase the number of samples used in the background coverage data set for XHMM (version 1.0), 220 samples were recruited, including the families of this project, sequenced with three different exome target kits (Agilent sureSelect V4, V5 and V6). The regions targeted by all three target kits were extracted using `bedtools multiinter` (version 2.24.0) and were merged if they were within a distance of 10bp. These regions were used as exome intervals. The XHMM analysis was performed following the recommended steps in the XHMM protocol (Fromer and Purcell, 2014).

From the resulting cohort level variant files, the candidate CNVs in a family were selected if the CNV is present in the case and the same variant should not be present in more than four other samples in the background data. Further, the CNVs are classified into *de novo* or inherited based on the absence or presence of the CNV in the parent. All the candidate variants were visualized and potential causal variants were selected for downstream analysis.

Ingenuity pathway analysis

Network analysis for validation of initially depicted candidate genes was performed using the knowledge database, Ingenuity Pathway Analysis (IPA). IPA core analysis was performed applying a merged list of genes known to be relevant for ENS development and to be HSCR disease-causing (Gui *et al.*, 2017; Lake and Heuckeroth, 2013; Luzón-Toro *et al.*, 2015b) as well as for re-filtered lists of candidate genes from sequenced trios of the pilot study (Master Thesis Volz, 2014) (tab. S3, Supplementary). In total, 15 networks were generated in the core analysis. Next, the four networks, which contained the candidate genes *ATP7A, SREBF1, ABCD1* and *PIAS2* were merged. Validated HSCR risk loci based on Luzon-Toro *et al.*, were manually added and connected to the merged network by the IPA software. Subsequently, indirect connections, which were not linked to any candidate gene product, were removed and networks were further modified as molecules were arranged corresponding to their subcellular localisation and non-connected factors were removed.

Genetic evaluation of candidate genes

Further HSCR patients carrying exonic variants in the candidate genes *SREBF1* and *PIAS2* were collected from already available sequencing data (whole genome and exome sequencing data) generated within the International HSCR consortium (Stanislas Lyonnet, http://www.erare.eu/financed-projects/hscr). Data analysis for *ATP7A* and *ABCD1* is ongoing.

Numbers of patients and controls used in the respective studies are summarized in the table below. Data from the following HSCR consortium members were taken into account: Prof. Dr. Robert Hofstra, Dr. Salud Borrego and Prof. Dr. Paul Kwong-Hang Tam. Phenotypic data of patients as well as genetic data from study specific control cohorts were accessible for the Spanish cohort only (tab. 6).

Table 6: HSCR patient cohorts from International HSCR Consortium groups
More detailed information for each cohort, which was taken into account for data analysis, is displayed. #: number

Cohort	**# of HSCR patients** (phenotype information)	**# of HSCR study controls** (origin, genotype information)
EUROPE		
The Netherlands	**38** (no phenotype information)	**212** (combined controls from EUROPE and SOUTH ASIA, no genotype information)
Spain	**56** (familial/sporadic; S-/ L-HSCR phenotypes)	**267** (Spanish controls from the Medical Genome Project, genotype data accessible via the Collaborative Spanish variant server (http://csvs.babelomics.org/))
SOUTH ASIA		
Hongkong	**443** (no phenotype information)	**493** (controls from other sequencing studies, no genotype information)

Lists of genetic variants found in HSCR patients were filtered in a second step in order to exclude common variants/SNPs (MAF >1%). First, all variants except for non-synonymous (missense mutations, stop loss and stop gain) were removed. Next, remaining variants were compared to variant data deposited in the gnomAD browser (population matched comparison). Variants were kept if MAF < 1%. To assess the putative functional relevance of the filtered variants, CADD scores were calculated using the CADD model GRCh37-v1.4 (https://cadd.gs.washington.edu/snv).

Statistics

All statistics were performed using the GraphPad Prism Software (version 5). The global significance level used was p=0.05.

An exploratory data analysis was carried out for quantitative real time (qRT) PCR expression data. For this purpose, mean expression levels for a specific gene and an individual time point were compared between the *mock control* clone and each of the gene-specific *KO* clones using a two-sided unpaired T-

Test (with Welch's correction in case of significantly different standard deviations of the two means compared). Corresponding to this analysis, statistical results are purely descriptive.

For functional *in vitro* assays targeting the cell proliferation, survival and migration capacity of the clones, mean values were compared between the *mock control* clone and each of the gene-specific *KO* clones using a one-way ANOVA. Results were Bonferroni corrected for multiple testing against the number of pairwise comparisons (*mock control* vs *RET KO/ ATP7A KO/ SREBF1 KO/ ABCD1 KO)*.

Candidate gene scoring

To evaluate the functional relevance and potential predisposing role of selected candidate genes for HSCR, a scoring system taking complementary data into account was established within a Bachelor Thesis project (Bachelor Thesis Dawid, 2019). It was aimed to summarize candidate specific evidences annotated in different databases or generated by *in silico* (IPA Network analysis) and *in vitro* (Transcriptomics assay) analyses in a single score. Data was classified into different categories (Gene expression, Gene information, Gene interaction, Gene constraint, Gene mutation) and ranked by a binary system. Classifications and cut offs for the respective categories are listed in tab. 7. Primary data, which was taken into account for risk evaluation, is partially displayed in sections 3.1 and 3.2. Hyperlinks for the databases are given in 2.1.15. More detailed information is provided in Dawid's Bachelor Thesis, 2019.

Table 7: Scoring system for risk evaluation of candidate genes

Categories, respective databases, software and experiments, which were taken into account for risk evaluation, are summarized. Detailed information about tissue, that was considered for gene expression evaluation is outlined below (see footnote), maximal scores for Σ_1 and Σ_2 are annotated in brackets

Category	**Gene expression**						**Gene information**			**Gene interaction**	**Gene constraint**		**Gene mutation**			
Database	GTEx portal, The Human Protein Atlas						Pubmed, Ensembl, OMIM, SFARI Gene, GWAS catalog, Genecards				gnomAD					
Software										IPA (2.2.1)			CADD SNV lookup, MutationTaster, PROVEAN			
Experiments	Microarray profiling (2.2.3)															
Classification/cut offs	mRNA level				Protein level		Function	Disease		Any connection to ENS/HSCR-related proteins	LoF oe value[8]		CADD	Mutation taster	PROVEAN	
	Human brain[1]	Human colon[2]	Human fetal colon[3]	Murine ENS tissue[4]	Human brain[5]	Human colon[6]	Neuron specific	Neurological disease	Inflammatory gut disorder (IBS/IBD)[7]		<0.35	**Σ_1**	>13	disease-causing	deleterious	**Σ_2**
Point	1	1	1	1	1	1	1	1	1	1	1	**(11)**	1	1	1	**(14)**
Σ/category	6						3			1	1	**11**	3			**14**

[1]Human Brain-Amygdala, Anteriorcingulate cortex, Caudate, Cerebellar hemisphere, Cerebellum, Cortex, Frontal cortex, Hippocampus, Hypothalamus, Nucleus accumbens, Putamen, Spinal cord, Substantia nigra
[2]Human Colon-Sigmoid, Transverse
[3]Microarray profiling of human fetal gastrointestinal cells from the colon (FGIC_C, ATCC), carried out by Dr. Stefanie Schmitteckert (2.2.3)
[4]Microarray profiling of murine tissue at E8.75, E9.5, E13.5 (2.2.3)
[5]Human Cerebral cortex, Hippocampus, Caudate, Cerebellum
[6]Human Colon, Rectum
[7]IBS: Irritable bowel syndrome, IBD: Inflammatory bowel disease
[8]LoF oe: Loss-of-function; observed/expected

2.2.2 *Cell biological methods*

All cells were cultivated at 37°C in a humidified atmosphere (95%) containing 5% CO_2.

Cell cultivation of HEK293TN cells

HEK293TN cells were cultivated in HEK293TN culture medium (HEK293TN CM) and passaged at 80-90% confluency. For passaging, media was aspirated and cells were washed once with pre-warmed 1X PBS. PBS was removed, pre-warmed 0.05% Trypsin-EDTA was added and cells were incubated at 37°C until having completely detached from the flask's surface. Trypsinization was stopped by the addition of pre-warmed HEK293TN CM. For further cultivation at 37°C, cell suspension was split and transferred into new cell culture flasks supplied with pre-warmed HEK293TN CM.

Cell cultivation and neuronal differentiation of SHSY5Y cells

SHSY5Y cells were cultivated in SHSY5Y culture medium (SHSY5Y CM) and passaged as previously described. Deviating from the outlined protocol, cell suspension was centrifuged at 1300 rounds per minute (rpm) for 3 minutes (min) before transferring it into new cell culture flasks.

To establish a neuronal differentiation protocol for WT SHSY5Y cells, modified versions of already published differentiation protocols were tested (Pilot experiment A) (Agholme *et al.*, 2010; Encinas *et al.*, 2000; Raghavan and Bitar, 2014). Further modifications to these setups were applied in a second pilot experiment (Pilot experiment B). Complete media compositions are given in 2.1.5. Three different coatings based on collagen I (Col I), Matrigel or collagen I, collagen IV and laminin (Col I, Col IV, laminin) were used. Details are described in the sections below. The experimental strategies are illustrated schematically in figure 3.

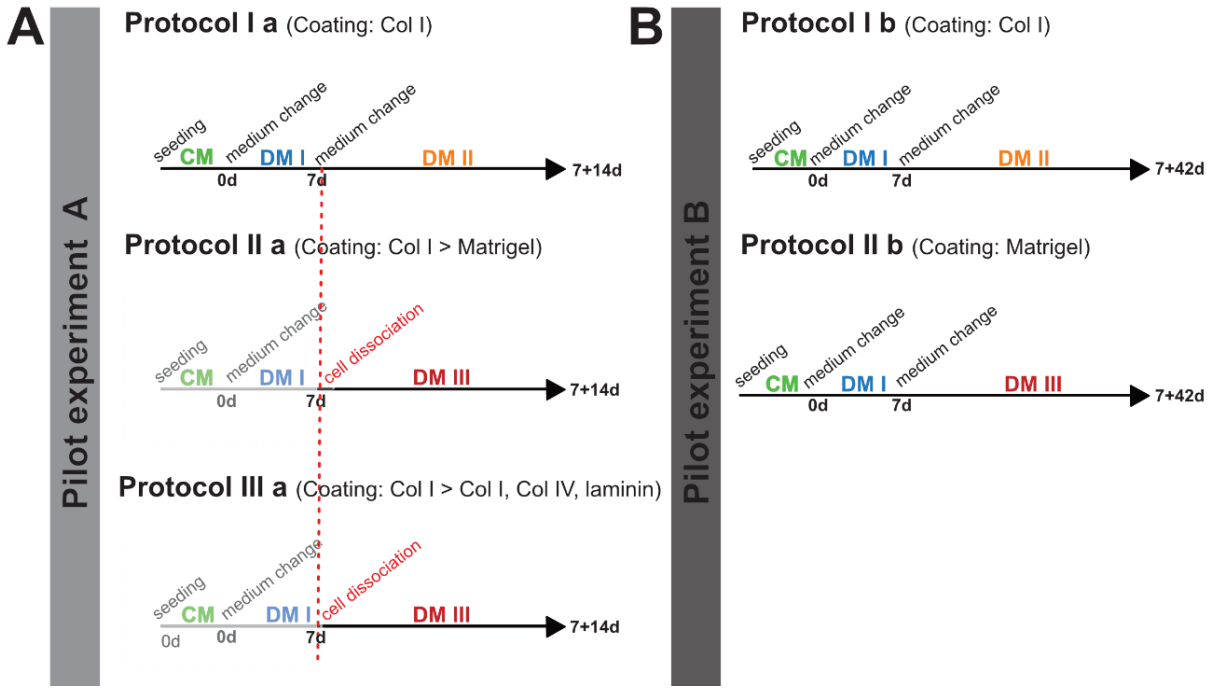


Figure 3: Establishment of a neuronal cell culture model system using SHSY5Y cells
A and B: Two pilot experiments (A, B) in SHSY5Y cells were carried out. Coatings, media, media changes, cell dissociation steps are outlined in relation to the respective time frame. Col I (Collagen I), Col IV (Collagen IV), CM (SHSY5Y culture medium), DMI I/DM II/ DMIII (SHSY5Y differentiation medium I/ II/ III), d (days)

Pilot experiment A

1x10^4 SHSY5Y cells/cm^2 were seeded on 0.05mg/mL collagen coated culture dishes in SHSY5Y CM. The day after, medium was aspirated and SHSY5Y differentiation medium I (SHSY5Y DM I) containing 10µM retinoic acid (RA) was applied. Medium was changed after 3-4 days (d) of cultivation. 7d after RA treatment, cells were further differentiated using three different experimental setups as described in the following:

Protocol Ia Cells were washed three times with DMEM high glucose and SHSY5Y DM II was applied (modified from Encinas *et al.*, 2000).

Protocol IIa Cells were washed once with pre-warmed 1X PBS and trypsinized as reported previously. 2x10^4 cells/cm^2 were plated in SHSY5Y DM III on Matrigel (1:3 in DMEM high glucose) coated dishes (modified from Agholme *et al.*, 2010).

Protocol IIIa: Cells were washed once with pre-warmed 1X PBS and trypsinized as reported previously. 2x10^4 cells/cm^2 were plated in SHSY5Y DM III on 800µg/mL Col I, 200µg/mL Col IV, 10µg/mL laminin, 1% FBS, 0.1% Pen/Strep (in DMEM high glucose, pH 7.6) coated dishes (modified from Agholme *et al.*, 2010 and Raghavan and Bitar, 2014).

Cells were differentiated for further 14d and harvested at specific time points (3d DM I, 7d DM I, 7+3d DM II/ DM III, 7+7d DM II/ DM III, 7+14d DM II/ DM III; n=1, each).

Pilot experiment B

Protocols Ia and IIa were further modified.

Protocol Ib: Cell seeding and differentiation were performed as previously described (Protocol Ia). Cultivation period was extended to 7+42d.

Protocol IIb: Experimental procedure was comparable to protocol Ib. Instead of Col I coated dishes, Matrigel (1:3 in DMEM) coated cell culture dishes were used. SHSY5Y DM III was used instead of SHSY5Y DM II.

Cells were harvested after 5d DM I, 7d DM I, 7+7d DM II/ DM III, 7+14d DM II/ DM III, 7+21d DM II/ DM III, 7+28d DM II/ DM III and 7+42d DM II/ DM III (n=2, each).

For differentiation of genome edited cell clones, differentiation protocol Ib was further modified: Cells were plated in SHSY5Y CM at a density of 2x10^4 cells/cm^2 on 0.05mg/mL Col I coated dishes. One day after seeding, medium was changed to SHSY5Y DM I containing RA which was changed after 3-4d of cultivation. 7d after RA treatment, cells were washed two times with DMEM high glucose. Cell differentiation in SHSY5Y DM II was performed for additional 28d (7+28d). Half of the media was changed every 4-7d.

Cell transfection

For evaluation of the candidate genes carrying the detected genetic variant, HEK293TN and SHSY5Y cells were transfected with respective overexpression constructs. For this purpose, coverslips were coated with Col I (0.05mg/mL) prior to cell seeding. 3.5x10^4 of HEK293TN and 1.2x10^5 of SHSY5Y cells

were seeded onto coated coverslips in 500µL of antibiotics free HEK293TN and SHSY5Y CM, respectively. Next day, HEK293TN cells were transfected using Lipofectamine 2000 following the manufacturer's instructions (500ng of plasmid DNA; ratio DNA/Lipofectamine 1:3). Medium was changed 4 hours (h) after transfection and cells were further cultivated for 48h before fixation. For transfection of SHSY5Y cells, the Magnetofectamine Kit (CombiMag/Lipofectamine 2000) was used. Transfection mixes were prepared as outlined in the manual and added dropwise to the cells (500ng of plasmid DNA; ratio DNA/Lipofectamine 1:4; ratio plasmid DNA/CombiMag 1:1). After 20min of incubation on a magnetic plate under standard conditions at 37°C, medium was changed. Cells were cultivated for further 48h before fixation.

CRISPR/Cas9 genome editing of SHSY5Y cells

Genome editing of SHSY5Y cells was performed as described in Master Thesis Mederer, 2015. In brief, $7.5x10^5$ cells were seeded per 6 well in a total volume of 2.5mL of SHSY5Y CM without the antibiotics penicillin and streptomycin. Next day, cells reached 80-90% confluency and were transfected with Lipofectamine 2000 and gene-specific sgRNAs cloned into the expression vector pSpCas9(BB)-2A-GFP. Transfection was carried out corresponding to the manufacturer's instructions (2.5µg of plasmid DNA; ratio DNA/Lipofectamine 1:3). For the generation of a *mock control* clone, 2.5µg of the empty cloning plasmid (pSpCas9(BB)-2A-GFP) were transfected. 4h after transfection, medium was changed to SHSY5Y CM and cells were cultured for further 48h at standard conditions.

For clonal isolation of genetically modified cells, fluorescent-activated cell sorting (FACS) at the FACS Core Facility (Dr. Volker Eckstein, University Hospital Heidelberg) was performed. Therefore, cells were harvested 48h after transfection corresponding to the aforementioned passaging procedure. To guarantee for single cell suspension, dissociated cells were resuspended with SHSY5Y CM and passed through a cell strainer (40µm). Next, cells were centrifuged at 1300rpm for 3min, supernatant was removed and the cell pellet was resuspended in 1–1.5mL ice-cold MACS buffer. Cell suspensions were kept on ice until FACS was performed. Single cells were sorted into single wells of up to four 96 well plates containing SHSY5Y CM. Cell clones were expanded for at least 21d before replica plates were prepared. To this end, cells were washed once with pre-warmed PBS and trypsinized as described previously. Trypsinization was stopped by cell resuspension in pre-warmed SHSY5Y CM and cell suspension was split into two 96 wells. Next day, medium was changed to remove remaining Trypsin-EDTA. Replicated 96 wells of individual clones were used for DNA extraction as described below.

The *mock control, RET KO* and *PIAS2_Ex6 KO* clones have already been generated in a previous study (Master Thesis Mederer, 2015). Gene-specific *KO* clones for *ATP7A, SREBF1* and *ABCD1* were generated in the context of this actual PhD project.

Boyden Chamber migration assay

Cell migration capacities of undifferentiated and differentiated genome edited cell clones were investigated by Boyden Chamber assays. In brief, 6.5mm transwell inserts (5µm pore size/ polycarbonate (PC) membrane) were coated with 0.05mg/mL Col I.

In the undifferentiated state, $1.2x10^5$ cells were seeded on coated inserts in serum-free media (DMEM high glucose). To measure the directed migratory behaviour of the cells, lower compartments were either

filled with SHSY5Y CM (FBS as chemoattractant; positive control) or DMEM high glucose only (no chemoattractant; negative control) (fig. 4 A).

To assess migration capacity in the differentiated state, cells were dissociated with Accutase 7d after initiation of neuronal differentiation. $1.5x10^5$ cells were seeded on Col I coated inserts as previously described. SHSY5Y DM II (BDNF as chemoattractant; positive control) or DMEM high glucose only (co chemoattractant; negative control) were added to lower compartments (fig. 4 A).

24h after incubation at standard conditions, cells from the inside of the inserts were removed using 1X PBS wetted cotton swabs. Migrated cells were fixed in ice-cold MeOH for 20min at room temperature (RT). After washing with 1X PBS, cells were counterstained with Hoechst (1:5000 in 1X PBS, 4min at RT) and mounted with Vectashield on a glass slide. For image analysis, membranes were imaged at five spots (20X magnification) as illustrated below using the Leica Automated Inverted Microscope DMI4000B (fig. 4 B). Nuclei were counted using the ImageJ software and the mean cell number/membrane was calculated. At least, three biological replicates were used for data analysis.

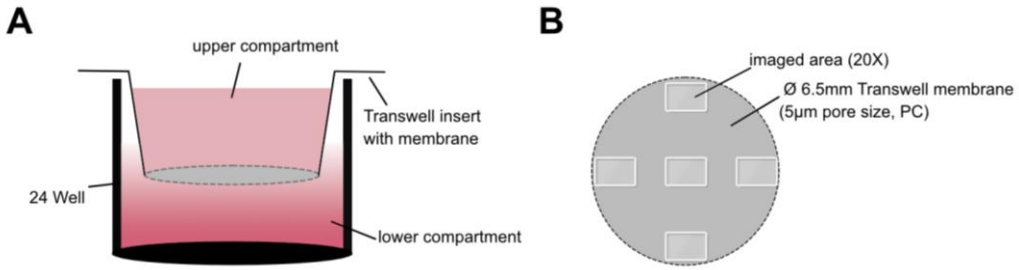


Figure 4: Boyden chamber migration assay
A: Schematic showing the experimental setup in a well of a 24 well plate. Upper and lower compartments are separated by a Transwell membrane. **B:** Imaged areas of the Transwell polycarbonate (PC) membrane are highlighted.

Cell proliferation assay

To assess the proliferative capacity of undifferentiated and differentiated genome edited cell clones, the Bromodeoxyuridine (BrdU) Incorporation Kit from Roche was used. For analysis of undifferentiated cells, $2.25x10^5$ cells were seeded on Col I (0.05mg/mL) coated coverslips. After 24h, medium was aspirated and BrdU labelling solution (10µM BrdU) diluted in SHSY5Y CM was added to the cells. After 10h of further cultivation under standard conditions, cells were washed and fixed corresponding to the manufacturer's instructions. For the analysis of differentiated cells, neuronal differentiation was initiated by 7d RA treatment as described earlier. Thereafter, the medium was changed to SHSY5Y DM II and the next day, BrdU labelling solution (10µM BrdU) diluted in SHSY5Y DM II was added, cells were cultivated for further 10h and processed as described previously for undifferentiated cells. Nuclei were counterstained with Hoechst (1:5000 in 1X PBS, 4min at RT). After mounting with Vectashield, cells were imaged with the Leica Automated Inverted Microscope DMI4000B. 10 fields of view (20X magnification) per cell clone were analysed using the ImageJ software. The mean percentage of BrdU positive cells/field of view was calculated. At least three biological replicates were used for data analysis.

Cell death detection assay

To determine the number of apoptotic cells, the *in situ* cell death detection Kit from Roche was used. **T**dT-mediated d**U**TP-X **n**ick **e**nd **labelling** (TUNEL) of apoptotic cells was performed as outlined in the protocol. For analysis of undifferentiated clones, $2.5x10^5$ cells were seeded cells on Col I (0.05mg/mL) coated cover slips and fixed in 4% PFA/ 1X PBS (20min at RT) after 24h. Differentiated cell clones after 7+7d of neuronal differentiation were investigated accordingly. Counterstain with Hoechst (1:5000 in 1X PBS, 4min at RT) was performed for undifferentiated and differentiated cell clones. Cells were mounted with Vectashield and imaged with the Leica Automated Inverted Microscope DMI4000B. 10 fields of view per cell clone (20X magnification) were used for image analyses (ImageJ software). The mean percentage of TUNEL positive cells/field of view was calculated from at least three biological replicates.

Preparation of cryosections

For preparation of whole mount cryosections, CD1 WT mice were sacrificed by CO_2 asphyxiation before the uterus was dissected. Embryos at stages E9.5, E10.5, E11.5 and E13.5 were prepared and collected in 1X PBS on ice. For preparation of adult tissue specimens, CD1 WT mice were sacrificed by CO_2 asphyxiation, part of the colon was dissected and rinsed multiple times with 1X PBS.

Fixation of whole mounts and gut tissue was carried out for 24h at 4°C in 4% PFA/ 1X PBS under gentle agitation. Next, tissues were washed twice with 1X PBS and incubated for 24h at 4°C in 30% sucrose (in DEPC-PBS) under gentle agitation. Finally, specimens were embedded in Tissue Freezing Medium, snap frozen in liquid nitrogen and stored at -80°C until cryosectioning with the cryostat-microtome CM3050S. Cryosections of 10µm thickness were prepared, placed on a glass slide and stored at -80°C until immunofluorescence (IF) staining.

IF staining of cryosectioned tissue

All antibody specific details are given in table 1 and 2.

For IF analyses, frozen cryosections were allowed to adapt to RT and then washed once with 1X PBS. Depending on the applied primary antibody, heat-induced antigen retrieval was performed for 25min in pre-warmed citrate buffer (pH 6) in a glass cuvette using a heat steamer. Sections were allowed to cool for 20min at RT before washing once with 1X PBS. After permeabilization for 7 min at RT using an antibody specific detergent (Triton X-100 or Tween-20) at a specific concentration, blocking was carried out for 1h at RT in 1% BSA/ 10% NGS/ 1X PBS supplemented with the respective antibody specific detergent. Primary antibodies diluted in blocking solution were incubated overnight (o/n) at 4°C in a humidified chamber. Next day, cryosections were washed three times with 1X PBS (each 5min) in a glass cuvette under gentle agitation. Secondary antibody dilutions were prepared in blocking solution and specimens were incubated for 1h at RT in the dark. After three washing steps with 1X PBS (each 5min), nuclear counterstain with Hoechst (1:5000 in 1X PBS) was performed for 4min in the dark. Next, sections were rinsed once with 1X PBS followed by a final washing step with H_2O. Stained slices were mounted with Vectashield and stored at 4°C until imaging using the Leica Automated Inverted Microscope DMI4000B.

For triple IF staining of cryopreserved tissue sections, slices were permeabilized, blocked and incubated with primary antibody solution (anti-SMA and the candidate specific antibody) o/n at 4°C. IF analyses of

these antigens were finalised as previously described. However, before nuclear counterstaining was applied, slices were blocked again and stained by anti-TUBB3 as outlined above. Next day, after secondary antibody incubation, cryosections were counterstained and mounted.

As negative control for the staining of cryoconserved tissue sections, no primary antibody incubation was carried out (fig. S2, Supplementary).

Preparation of formalin-fixed paraffin embedded sections

Fetal and adult formalin-fixed paraffin embedded (FFPE) tissue sections were kindly prepared by Dr. Cornelia Thöni and Jutta Scheuerer (Institute of Pathology, University Hospital Heidelberg, research team of Dr. Felix Lasitschka). Human fetal gut tissue originates from a fetus acardius amorphous (25th week of gestation (wog)) while adult colon samples originated from a non-affected part of the colon of a patient suffering from Inflammatory bowel disease (IBD).

Immunohistochemical staining of human tissue slices

Primary antibody specific details are given in table 1.

Immunohistochemical (IHC) staining of fetal and adult FFPE tissue sections were kindly performed by Dr. Cornelia Thöni and Jutta Scheuerer (Institute of Pathology, University Hospital Heidelberg, research team of Dr. Felix Lasitschka). In brief, FFPE samples of 3µm thickness were deparaffinized and rehydrated. Heat induced antigen retrieval was performed in 10mM EDTA (pH 9) for 5min. Slices were blocked in antibody diluent solution (Thermo Fisher Scientific) supplemented with 15% NGS (Vector Laboratories). Specimens were incubated with primary antibodies diluted in blocking solution o/n at 4°C. Next day, slides were washed three times with TBST (0.5% Tween-20), 10min each. Peroxidase-conjugated AffiniPure Goat Anti-Rabbit (Jackson ImmunoResearch) was diluted in blocking solution and secondary antibody incubation was carried out for 1h at RT. Again, slides were washed three times with TBST (10min each) before 3,3′-Diaminobenzidine (DAB) solution (ImmPACT DAB Peroxidase Substrate Kit) was added onto the slices. DAB reaction was stopped when signal intensity of choice was reached by washing with ddH_2O. Slides were counterstained with Hematoxylin for 1min. After washing with ddH_2O, slides were differentiated in running tap water and mounted with Aquatex.

Staining was imaged using the Olympus Light Microscope BX53 and the Olympus SC30 camera (Institute of Pathology, University Hospital Heidelberg).

As negative control for the staining of FFPE tissue sections, no primary antibody incubation was carried out (fig. S3, Supplementary).

IF staining of cells

Antibody specific details are given in table 1 and 2.

For cell fixation, respective media was removed, and cells were washed once with 1X PBS. Fixation was carried using 4% PFA (in PBS). After 20min of incubation at RT, fixative was removed, and cells were washed three times with 1X PBS. Coverslips were stored in 1X PBS at 4°C until IF staining were carried out. Differentiated genome edited cell clones were cultivated on 3.5cm ibidi imaging dishes until fixation with 4% PFA/ 1X PBS which was performed as previously described.

IF staining of fixed cells was performed corresponding to the following protocol: for staining of nuclear antigens, fixed cells were permeabilized with 0.1% Triton (in 1X PBS) for 7min at RT prior to blocking for 1h at RT in 1% BSA/ 10% NGS/ 1X PBS. After blocking, cells were incubated with primary antibodies diluted in blocking solution in a humidified chamber for 1h at RT. Next, cells were washed three times with 1X PBS and incubated for 1h with secondary antibody dilution (in blocking reagent) at RT in a humidified chamber in the dark. Again, cells were washed three times with 1X PBS before nuclear counterstaining with Hoechst (1:5000 in 1X PBS) was performed for 4min under light protection. After two washing steps with 1X PBS and a final washing step with H_2O, cells were prepared for fluorescent microscopy using the Leica Automated Inverted Microscope DMI4000B. Therefore, coverslips were mounted with Vectashield on a glass slide while cells on ibidi dishes were mounted with AquaPolymount.

2.2.3 *Molecular biological methods*

Total RNA isolation

Total RNA extraction was performed with TRIzol reagent either following the manufacturer's instructions or combined with the RNAqueous-Micro Total RNA Isolation Kit protocol. To this end, TRIzol samples were thawed on ice and 200µL of chloroform/mL were added. The samples were mixed vigorously for 15 seconds (sec). After incubation for 5min at RT, specimens were centrifuged for 15min at 4°C at 13000rpm. The upper phase was transferred into a new tube and 0.5volumes of EtOH were added. After vigorous mixing for 5sec, 150µl of the sample was transferred onto the column of the RNAqueous-Micro Total RNA Isolation Kit. All further steps were carried out corresponding to the manual.
Before further usage, the quality and quantity of the RNA samples were assessed by a spectrophotometer from DeNovix.

Transcriptomics assay (Clariom S array mouse)

RNA samples from different murine embryonic tissue were used for transcriptomics analyses at the DKFZ Genomics and Proteomics Core Facility using the Clariom S assay (mouse) (Thermo Fisher Scientific). Regions containing pre-migratory vagal NCCs at stage E8.75 and E9.5 as well as the embryonic gut at E13.5 were dissected from CD1 WT murine embryos.
Dissected embryonic tissues from each litter were pooled. Three litters were analysed for each stage. Total RNA was extracted using TRIzol reagent according to the manual. In addition to the RNA analysis by the DeNovix system, the RNA amount was quantified and checked for integrity by the Qubit 2.0 Fluorometer and the Agilent 2100 Bioanalyzer System according to the manufacturer's instructions, respectively. 500ng of total RNA (RIN from 9.8-10) of each preparation were submitted for microarray profiling at the DKFZ Genomics and Proteomics Core Facility (Dr. Melanie Bewerunge-Hudler).
For gene expression profiling, biotinylated antisense cDNA was subsequently prepared according to the standard labelling protocol with the GeneChip WT Plus Reagent Kit and the GeneChip Hybridization, Wash and Stain Kit (both from Thermo Fisher Scientific). Afterwards, the hybridization on the chip was performed on a GeneChip Hybridization oven 640, then dyed in the GeneChip Fluidics Station 450 and thereafter analysed by a GeneChip Scanner 3000. All of the equipment used was from the Affymetrix. Data analysis was kindly performed by Dr. Carolina De La Torre from the Medical Faculty Mannheim. In brief, a Custom CDF Version 22 with ENTREZ based gene definitions was used to annotate the

arrays (Dai *et al.*, 2005). The raw fluorescence intensity values were normalized applying quantile normalization and robust multi-array (RMA) background correction.

The expression profile of undifferentiated human fetal gastrointestinal colon cells (FGIC_C, ATCC) was assessed in the context of another project by microarray profiling according to the previously presented pipeline (Dr. Stefanie Schmitteckert, unpublished data). Respective data was used for the establishment of the scoring system (Bachelor Thesis Dawid, 2019).

cDNA synthesis

cDNA synthesis was performed with the SuperScript III First-Strand Synthesis kit. Deviating from the manual, 50pmol Oligo dT and 50ng random hexamer primers were used. For reverse transcription of differentiated genome edited cell clones, 800ng of total RNA were used per sample. For conventional PCR, cDNA was diluted 1:10, whereas for quantitative realtime (qRT) PCR cDNA was diluted 1:5.

Conventional PCR

Conventional PCR analyses were performed with the HotStarTaq DNA Polymerase PCR Kit or the Q5 High fidelity DNA Polymerase Kit either using cDNA or genomic DNA as template. Thermal cycling was carried out in Mastercycler pro – vapo protect. Depending on the amplicon size, PCR products were separated by gel electrophoreses using 1%, 1.5% or 2% TAE agarose gels, respectively. *GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (GAPDH)* was used as control to assess RNA integrity and RT quality. For Sanger sequencing, PCR product purification was performed using the QIAquick PCR Purification Kit according to the instructions given in the manual. Sanger sequencing was carried out by GATC, EUROFINS or GENEWIZ.

qRT PCR

Quantitative gene expression analysis of *in vitro* differentiated gene-specific *KO* cell clones was performed using the SYBR Green Lo-ROX Fast Mix from Bioline. PCR conditions are shown in tab. S7 (Supplementary). A set of different markers was used for characterization of differentiating cells. Oligonucleotide specific sequences are given in table 3. Respective standard curves were generated by using cDNA of human whole brain or SHSY5Y 7d DM I. qRT PCR analysis for each marker and condition was performed as duplicate, while up to three biological replicates were used for each time point. Relative mRNA expression levels were calculated by the relative standard curve method and via normalization to the expression of the two reference genes *(SUCCINATE DEHYDROGENASE COMPLEX FLAVOPROTEIN SUBUNIT A (SDHA)* and *HYPOXANTHINE PHOSPHORIBOSYLTRANSFERASE 1) HPRT1* for WT SHSY5Y, and *SDHA* and *GAPDH* for all edited cell clones and the *mock control* clone.

Total Protein isolation

Undifferentiated cells were harvested by trypsinization as previously described. Differentiated and transfected cells were washed once with 1X PBS before scraping in 1X PBS on ice. Cell suspension was transferred into a tube and cells were pelleted by centrifugation for 3min at 1300rpm. Cell pellets of undifferentiated, differentiated as well as transfected cells were resuspended in the appropriated amount of RIPA buffer complemented with proteinase and phosphatase inhibitor cocktail.

For protein isolation of murine tissue, embryonic and adult GI and brain tissues were dissected and rinsed with sterile 1X PBS before resuspension in RIPA buffer complemented with proteinase and phosphatase inhibitor cocktail. Adult murine specimens were homogenized on ice using the polytron tissue homogenizer. Embryonic tissues were minced using a pistil.
Cell lysis was performed for 45min at 4°C under rotation. After centrifugation for 15min at 13000rpm at 4°C, the supernatant was transferred into a new tube and used for protein quantification with the Pierce BCA Protein Assay Kit.

Protein gel electrophoresis and Western blot analysis

8 or 10% SDS-polyacrylamid Laemmli gels were casted in Novex Gel cassettes and used for Western blot analyses of anti-RET, anti-SREBF1 and anti-PIAS2 specific investigations. 5-60µg of total protein per sample were prepared with 5X sample loading buffer and denatured for 5min at 95°C. Samples were loaded onto the gel and gelelectrophoresis in 1X Laemmli running buffer was carried out at 150 Volt (V) for 1.5-2h at 4°C.
Precasted NuPAGE 4-12% Bis-Tris gradient gels and the NuPAGE Bis-Tris Gel Protocol were used for protein gel electrophoresis of anti-ATP7A and anti-ABCD1 specific investigations. 5-60µg of total protein per sample were prepared with NuPAGE 4X sample loading buffer and denatured for 10min at 75°C. Gel electrophoresis in 1X MOPS running buffer was carried out at 200V for 1-1.5h at 4°C.
For protein transfer, a PVDF membrane was used. To this end, the membrane was wetted with MeOH before preparing the blotting sandwich as indicated below:

- two blotting pads soaked with transfer buffer
- two Whatman papers soaked with transfer buffer
- gel
- PVDF membrane
- two Whatman papers soaked with transfer buffer
- two blotting pads soaked with transfer buffer

Blots were run at 150mA for 2.5h at RT. Blotting for anti-ATP7A and anti-ABCD1 was carried out at 25V for 75min at RT using the NuPAGE Transfer buffer.
After protein transfer, the membranes were blocked in Odyssey blocking reagent (PBS) for 1h at RT. Primary antibodies were diluted in Odyssey blocking reagent (PBS) + 0.2% Tween-20 and incubation was carried out o/n at 4°C under gentle agitation. Next day, membranes were washed three times with 1X TBST for 5min each. Secondary antibodies were diluted in Odyssey blocking reagent (PBS) + 0.2% Tween-20. Incubation was carried out for 1h at RT in the dark under gentle agitation. Before imaging with the Odyssey Infrared Imaging System was performed, membranes were washed three times for 5min with 1X TBST and once with 1X TBS.
Respective antibody dilutions are listed in table 1 and 2.

Cloning of expression constructs

For cloning of sgRNA expression constructs into the pSpCas9(BB)-2A-GFP vector backbone, *BbsI* digested plasmid DNA was ligated with annealed and phosphorylated sgRNA oligonucleotides. Ligation was performed corresponding to the manual's instruction of the T4 DNA ligase Kit. Details are given in the Supplementary (tab. S6, fig. S8).

For cloning of gene-specific overexpression constructs, the pcDNA3.1-2xFLAG vector backbone was used. Inserts were generated using the Q5 High fidelity DNA Polymerase Kit. SHSY5Y cDNA served as template. Respective oligonucleotides are shown in table 3. PCR products were purified using the QIAquick PCR purification kit. After quantification with the DeNovix spectrophotometer, restriction digestion using FastDigest restriction enzymes was performed according to the manufacturer's instructions. Restriction enzymes used for gene-specific overexpression construct generations are listed below.

Insert	**Restriction enzymes used**
RET_FL	*SrfI*, *XbaI*
SREBF1b_FL	*BamHI*, *XbaI*

Digested inserts and plasmids were loaded onto a 1% agarose gel, respective gel bands cut out and the DNA isolated with the QIAquick Gel extraction Kit and quantified with the DeNovix spectrophotometer. 10-50ng of linearized plasmid DNA were ligated with respective excess amounts of insert DNA using the T4 DNA ligase corresponding to the instructions given in the protocol. Vector maps are displayed in the supplementary (fig. S9 and S10).

In case of compound heterozygous events, allele specific genomic modifications caused by CRISR/Cas9 genome editing in individual alleles were separated using the pSTBlue-1 AccepTor Vector Kit. All experimental steps were carried out corresponding to the instructions given in the manual. NovaBlue Singles Competent cells (>1.5x10^8 colony forming units (CFU)/µg plasmid DNA) were used for bacterial transformation following the manufacturer's instructions. To allow blue white screening, ampicillin containing LB agar plates were coated with IPTG/XGal before plating of the transformed bacterial cell suspension.

Patient-specific mutations were introduced into the respective overexpression constructs using the QuickChange Lightning Site-Directed Mutagenesis Kit following the manufacturer's instructions. Deviating from the manual, 1µl of DMSO was added to the *ATP7A* and *ABCD1* mutant strand synthesis reactions. Agilent Technologies XL10-Gold Ultracompetent Cells (≥5x10^9 colony forming units/µg plasmid DNA) were used for bacterial transformation. Instructions of the manual were followed.

If not otherwise stated, chemo-competent *DH5α E.Coli* cells (>1x10^6 CFU/µg plasmid DNA) were used for transformation. Bacterial cells were thawed on ice for 5min and 1µL of the ligation reaction was added. Cells were shortly mixed and incubated for 8min on ice before 250µL of pre-warmed SOC medium were added. For bacterial outgrowth, cells were incubated up to 60min at 37°C under constant agitation (350rpm). LB agar plates containing 50µg/mL ampicillin or 30µg/mL kanamycin were pre-

warmed at 37°C before up to 100µL of the transformed bacterial cell suspension were plated. Plates were incubated o/n at 37°C.

Plasmid preparation from bacteria

To sequence validate bacterial clones, 5mL of LB medium containing the respective antibiotics were inoculated with a single colony. After o/n incubation at 37°C under constant shaking (1000rpm), the GeneJET Plasmid MiniPrep Kit was used for plasmid preparation. Sanger sequencing was carried out at GATC, EUROFINS or GENEWIZ.

For midi plasmid preparations 50-100mL LB medium containing the respective antibiotics were inoculated with up to 20µL of an o/n mini prep culture. After bacterial outgrowing at 37°C o/n, the ZymoPURE II plasmid prep Kit was used for plasmid preparation.

Glycerol stocks of bacterial cultures were prepared by mixing 300µL of glycerol with 700µL of an o/n bacterial culture. After snap-freezing, glycerol stocks were stored at -80°C.

Genomic DNA extraction and genotyping

To detect micro indels in the CRISPR/Cas9 targeted genomic regions, Sanger sequencing of a PCR product spanning the region of interest was performed. To this end, replicated SHSY5Y cell clones were trypsinized as previously described and cell suspension was pelleted at 2000rpm for 4min. Genomic DNA extraction was performed using the DirectPCR Lysis-Reagent Cell following the manufacturer's instructions. Oligonucleotides used for genotyping are listed in table 3. The respective PCR product was purified using the QIAQuick PCR Purification Kit and sequenced at GATC or EUROFINS.

3 RESULTS

3.1 Evaluation of WES data and genotyping for non-coding HSCR SNPs

In two investigated families, *ATPase COPPER TRANSPORTING ALPHA* (***ATP7A***) and *STEROL REGULATORY ELEMENT BINDING TRANSCRIPTION FACTOR* (***SREBF1***) as well as *ABC SUBFAMILY MEMBER 1* (***ABCD1***) and *PROTEIN INHIBITOR OF ACTIVATED STAT 2* (***PIAS2***), represented most promising HSCR candidate genes. Bioinformatic data revision (performed by Dr. Nagarajan Paramasivam, Theoretical Bioinformatics Division, DKFZ) on the basis of current database states, confirmed this gene selection and reduced the lists of filtered candidate genes from 19 to 11 candidate genes for family I and from 20 to 15 candidate genes for family II (tab. S1 and S2, Supplementary). After variant calling, filtered WES data of both patients were as well checked for the presence of rare variants residing in ENS-relevant genes (tab. S3, Supplementary). Patient I harboured heterozygous variants in three ENS/HSCR-associated genes: *ECE1* (rs141146885), *SERPIN FAMILY I MEMBER 1* (*SERPINI1*) (rs61750375) and *SUPPRESSOR OF FUSED HOMOLOG* (*SUFU*) (rs34135067) while for patient II no variants in any ENS-relevant genes could be identified (fig. 5).

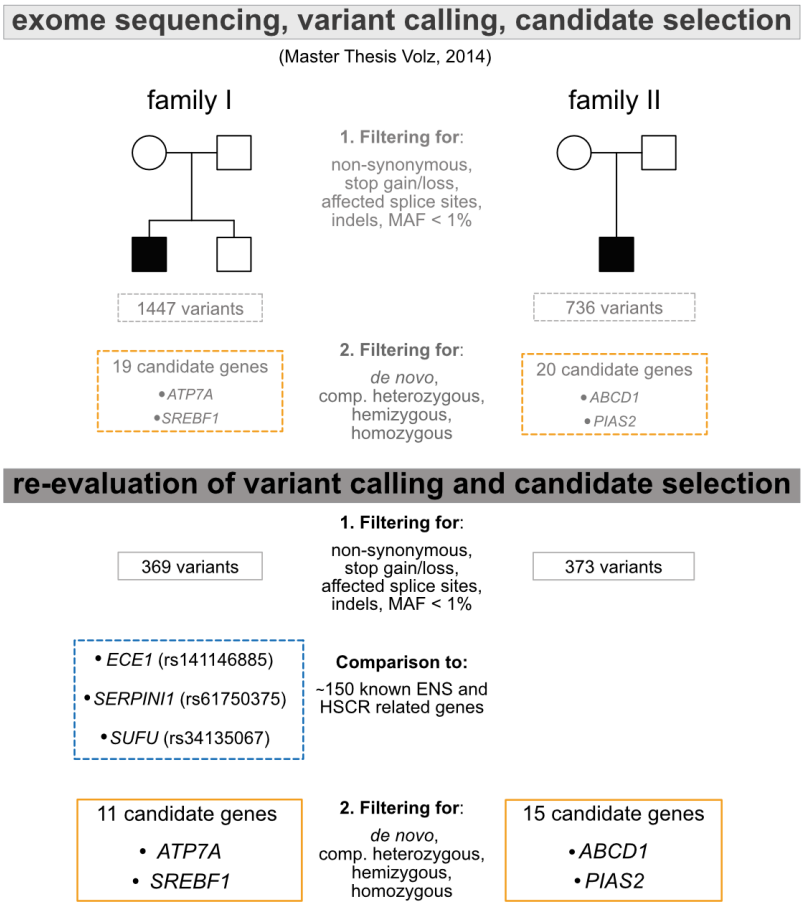


Figure 5: Schematic overview on the bioinformatics filtering of WES data
Variant calling and candidate selection of a previous study were re-evaluated and initial gene selection could be confirmed. MAF (Minor allele frequency), comp. (compound), indel (insertion/deletion)

Additionally, WES data was used for CNV analyses in both families (performed by Dr. Nagarajan Paramasivam, Theoretical Bioinformatics Division, DKFZ). These investigations led to the identification of several deleted and duplicated chromosomal regions which are summarized in fig. 6 A. For example, patient I carried an inherited, 30kb deletion within the *PROTO-CADHERIN GAMMA SUBFAMILY (PCDHG)* gene cluster affecting *PCDHGA1, PCDHGA2, PCDHGA3, PCDHGA4, PCDHGB1* and *PCDHGB2* (fig. 6 A'). Patient II carried a 9.5kb deletion on chromosome 8 affecting the *CLN8 TRANSMEMBRANE ER AND ERGIC PROTEIN (CLN8)* and a 60kb duplication in the *GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE 3 (AGPAT9)* residing on chromosome 4 (fig. 6 A'').

Except for genetic alterations in the coding part of the genome, all individuals of family I and II were genotyped for four non-coding HSCR risk SNPs located in *REARRANGED DURING TRANSFECTION (RET)* (rs2435357, rs7069590, rs2506030) and the *CLASS 3 SEMAPHORIN (SEMA3)* cluster (rs11766001). All genotype information on the above-mentioned variants is summarized in fig. 6 B.

Family I **Family II**

A **CNV analysis**

A'

	Deletion	Duplication
Mother	Chr 5:140725974-140755213; 30kb *(PCDHGA1/2/3/4,PCDHGB1/2)*	
Father		Chr. 6: 31371272-31380270; 9kb *(MICA)*, Chr. 12: 100431389-100466660; 35kb *(UHRF1BP1L)*
Patient	**Chr 5:140725974-140755213; 30kb *(PCDHGA1/2/3/4,PCDHGB1/2)***	**Chr 6: 31371272-31380270; 9kb *(MICA)*, Chr. 12: 100431389-100466660;35kb *(UHRF1BP1L)***
Brother		Chr. 6: 31371272-31380270; 9kb *(MICA)*, Chr. 12: 100431389-100466660; 35kb *(UHRF1BP1L)*

A''

	Deletion	Duplication
Mother		
Father	Chr. 8: 1719243-1728703; 9.5kb *(CLN8)*	Chr. 4: 84465639-84525975; 60kb *(AGPAT9)*
Patient	**Chr. 8: 1719243-1728703; 9.5kb *(CLN8)***	**Chr. 4: 84465639-84525975; 60kb *(AGPAT9)***

B **Genotyping non-coding HSCR risk SNPs**

B'

	RET			*SEMA3A*
	rs2435357 C>T	rs7069590 T>C	rs2506030 A>G	rs11766001 A>C
Mother	T/T	T/T	A/A	A/C
Father	C/C	T/T	A/A	A/A
Patient	**C/T**	T/T	**A/A**	**A/C**
Brother	C/T	T/T	A/A	A/C

B''

	RET			*SEMA3A*
	rs2435357 C>T	rs7069590 T>C	rs2506030 A>G	rs11766001 A>C
Mother	C/T	T/T	G/G	A/C
Father	C/T	T/T	A/G	A/A
Patient	**T/T**	**T/T**	**G/G**	**A/A**

Figure 6: Results for CNV analysis and genotyping for non-coding HSCR risk SNPs

A: Deleted and duplicated chromosomal regions detected in the individuals of the two families are shown. Except for the chromosomal position (based on GRCh37/hg19) and the size (in kb), the affected gene is annotated in brackets (Family I (A'): *PCDHGA1/2/3/4, PCDHGB1/2:* members of the *PROTO-CADHERIN GAMMA SUBFAMILY A/B; MICA: MHC CLASS I POLYPEPTIDE-RELATED SEQUENCE A; UHRF1BP1L: UHRF1 BINDING PROTEIN 1 LIKE;* Family II (A''): *CLN8: CLN8 TRANSMEMBRANE ER AND ERGIC PROTEIN; AGPAT9: GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE 3*). **B:** Genotype information of all four non-coding HSCR risk SNPs are listed for each individual. Risk alleles are highlighted in red.

Genetic data from all individuals of the two families is summarized in fig. S1 (Supplementary).

3.2 Candidate gene validation

3.2.1 *IPA Network analysis*

To put the detected candidates into biological context, an Ingenuity Pathway Analysis (IPA) was performed and revealed multiple direct and indirect protein interactions between depicted candidates and genes relevant in ENS formation or proven to be causative for HSCR (fig. 7). ATP7A displayed an indirect connection to the LOW-DENSITY LIPOPROTEIN (LDL) complex to which RETINOL BINDING PROTEIN 4 (RBP4) was directly connected. RBP4 was one of the ENS-relevant factors uploaded for network analysis. The very-long chain fatty acid (VLCFA) transporter ABCD1 was found to be directly connected to the HSCR risk factor DNA METHYLTRANSFERASE 3 BETA (DNMT3B). For the candidates SREBF1 and PIAS2, several direct protein interactions were assessed using the IPA knowledge base. Amongst others, both candidates showed a direct connection to the ENS-relevant factor NEURON NAVIGATOR 2 (NAV2). In addition, SREBF1 was identified to directly interact with the T-BOX3 transcription factor (TBX3) while for PIAS2 a direct connection to the PHOSPATASE AND TENSIN HOMOLOG (PTEN) was elucidated. Both candidate interactions partners are known regulators of ENS development.

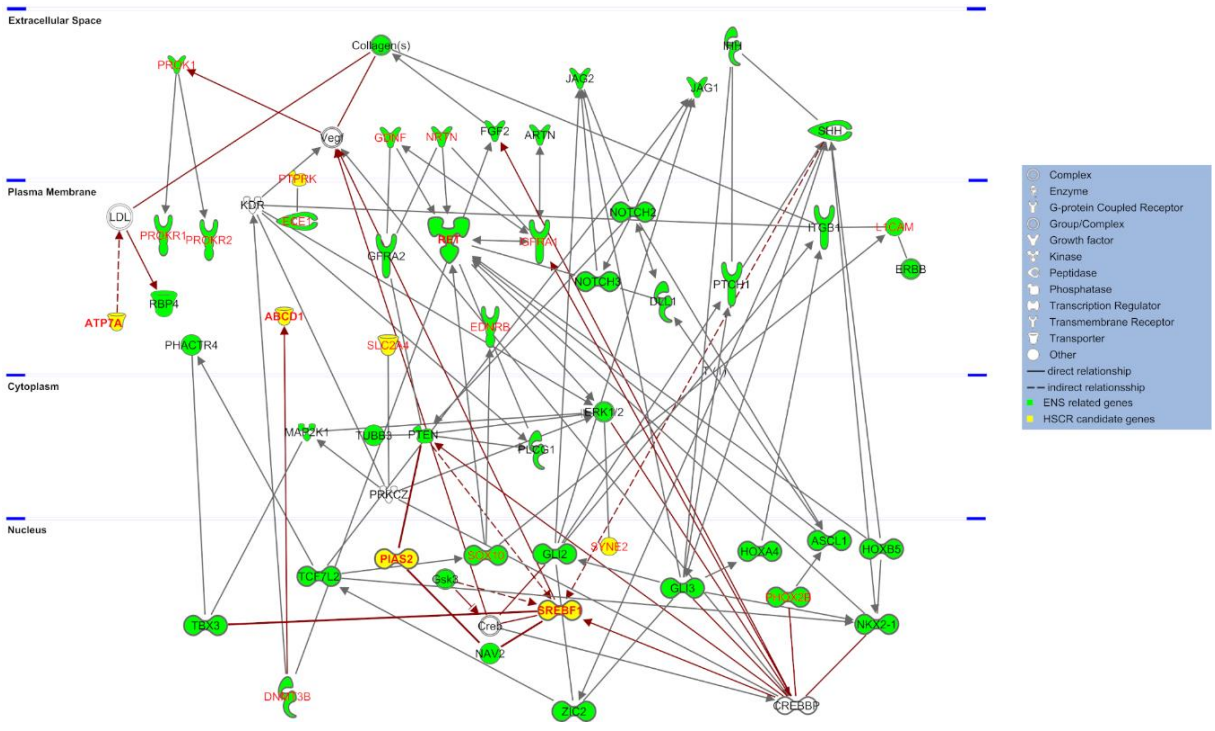


Figure 7: Network generated based on IPA

After core analysis using the Ingenuity Pathway Analysis (IPA) knowledge base, networks (4/15) containing the four candidate genes of interest were merged. Additionally, gene products of validated HSCR risk loci were added and connected to the network. Indirect interactions were only kept if they were linked to candidate gene products. Proteins known to be relevant for ENS development are marked in green. In addition, validated HSCR risk factors are shown in red. Candidates are colored in yellow. Interactions between candidates and ENS/HSCR relevant factors are highlighted in red. All proteins are arranged corresponding to their subcellular localization.

3.2.2 *Genetic evaluation of candidate genes*

To confirm the genetic relevance of the depicted candidate genes, additional mutation carriers needed to be identified. For this purpose, data from the International HSCR consortium was used. This genetic evaluation was already performed for two of the four selected candidate genes (*SREBF1* and *PIAS2)* in the context of this project while *ATP7A* and *ABCD1* will be investigated accordingly as soon as respective data is available.

Filtering of genetic variants within the respective patient cohorts (from Europe and South Asia) led to the identification of 11 additional HSCR patients carrying novel mutations within *SREBF1* (fig. 8 A and A') and 6 additional HSCR cases presenting with novel mutations within *PIAS2* (fig. 8 B and B'). The identified initial variants of *SREBF1* and *PIAS2* were not found in a second carrier. To determine if these filtered variants might have a functional relevance, CADD scores were taken into account, which showed that all filtered variants displayed scores >13. The corresponding data is summarized in fig. 8.

A

SREBF1

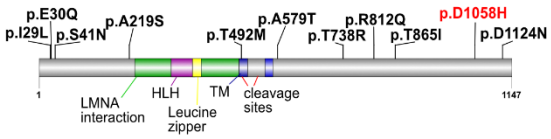
Variant		CADD	# of carriers	gnomAD comparison (allele count/ allele number)	Patient Cohort
Nucleotide change	Aminoacid change				
c.A85C	p.I29L	23.8	1	n. a.	EUR
c.C2213G	p.T738R	25.4	1	n. a.	
c.G2435	p.R812Q	18.62	1	259/48276 (0.005365)	
c.G3370A	p.D1124N	33	1	n. a.	
c.G88C	p.E30Q	29.3	1	n. a.	SA
c.G122A	p.S41N	15.83	1	n. a.	
c.G655T	p.A219S	15.64	2	n. a.	
c.C1475T	p.T492M	20.8	1	0/14960	
c.G1735A	p.A579T	23.4	1	n. a.	
c.C2594T	p.T865I	22.7	1	n. a.	

B

PIAS2

Variant		CADD	# of carriers	gnomAD comparison (allele count/ allele number)	Patient Cohort
Nucleotide change	Aminoacid change				
c.A236G	p.K79R	16.05	2	0/42766	EUR
c.C946T	p.P316S	23	1	0/10308	SA
c.C1522T	p.Q508*	44	1	n. a.	
c.A1627G	p.S543G	22	1	n. a.	
c.A1777G	p.T593A	16.65	1	0/15686	

A'



B'

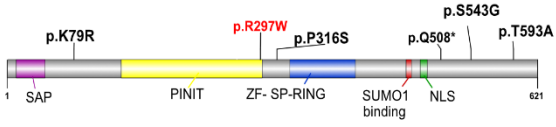


Figure 8: Results of genetic evaluation of candidate genes

A and A': Filtered rare variants residing in *SREBF1* are listed in a table format (A) and are schematically shown corresponding to their localisation on protein level (A'). **B and B':** Filtered rare variants residing in *PIAS2* are listed in a table (B) and are schematically shown on protein level (B'). Variants were identified in additional HSCR patients who have been analysed by different groups from the International HSCR consortium (http://www.erare.eu/financed-projects/hscr). For filtering, population matched control data deposited in the gnomAD browser were used (EUR (European), SA (South Asian)). CADD scores were calculated using the CADD model GRCh37-v1.4 (https://cadd.gs.washington.edu/snv).Nucleotide and aminoacid positions refer to SREBF1b (ENST00000261646.5) or PIAS2ß (ENST00000585916.1) (A, B). The gene-specific variants identified in the index cases of this project are highlighted in red (A', B'). Protein domains were annotated based on database entries (Pfam, NCBI, Uniprot). TM (Transmembrane), HLH (Helix loop Helix), LMNA (LAMININ A/C), ABC (ATP binding cassette), SAP (scaffold attachment factor-A/B, acinus and PIAS), PINIT (Pro-Ile-Asn-Ile-Thr), ZF-SP-RING (Zinc Finger – Really Interesting New Gene), SUMO (small Ubiquitin-related modifier), NLS (Nuclear localisation signal)

3.2.3 *Expression analyses on mRNA level*

Transcriptional profiles of ENS-relevant tissues were assessed by microarray analysis. Regions with pre-migratory vagal NCCs at E8.75 and E9.5 as well as the developing gut at E13.5 were dissected and analysed (fig. 9 A).

Based on hierarchical clustering of the mean values of the data distribution (overall mean expression), the three specimens could be distinguished (data analysis performed by Dr. Carolina de la Torre, Center of Medical Research, Mannheim). Correlation coefficients for comparisons between E8.75 and E9.5 were higher than respective coefficients for E13.5 and E9.5 or E8.75 (fig. 9 B).

log2 mRNA expression values of known HSCR risk loci ranged between 4-12 with *Neuregulin 3 (Nrg3)* showing lowest and *Transcription factor 4 (Tcf4)* displaying highest expression levels in general (fig. 9 C). The most prominent HSCR risk genes *Ret, Glial-derived neurotrophic factor (Gdnf), Gdnf family receptor alpha 1 (Gfra1), Endothelin receptor type B (Ednrb), Endothelin 3 (Edn3), Ece1, SRY box 10 (Sox10)* and *Paired like-homeobox 2b (Phox2b)* were expressed at intermediate to high levels and

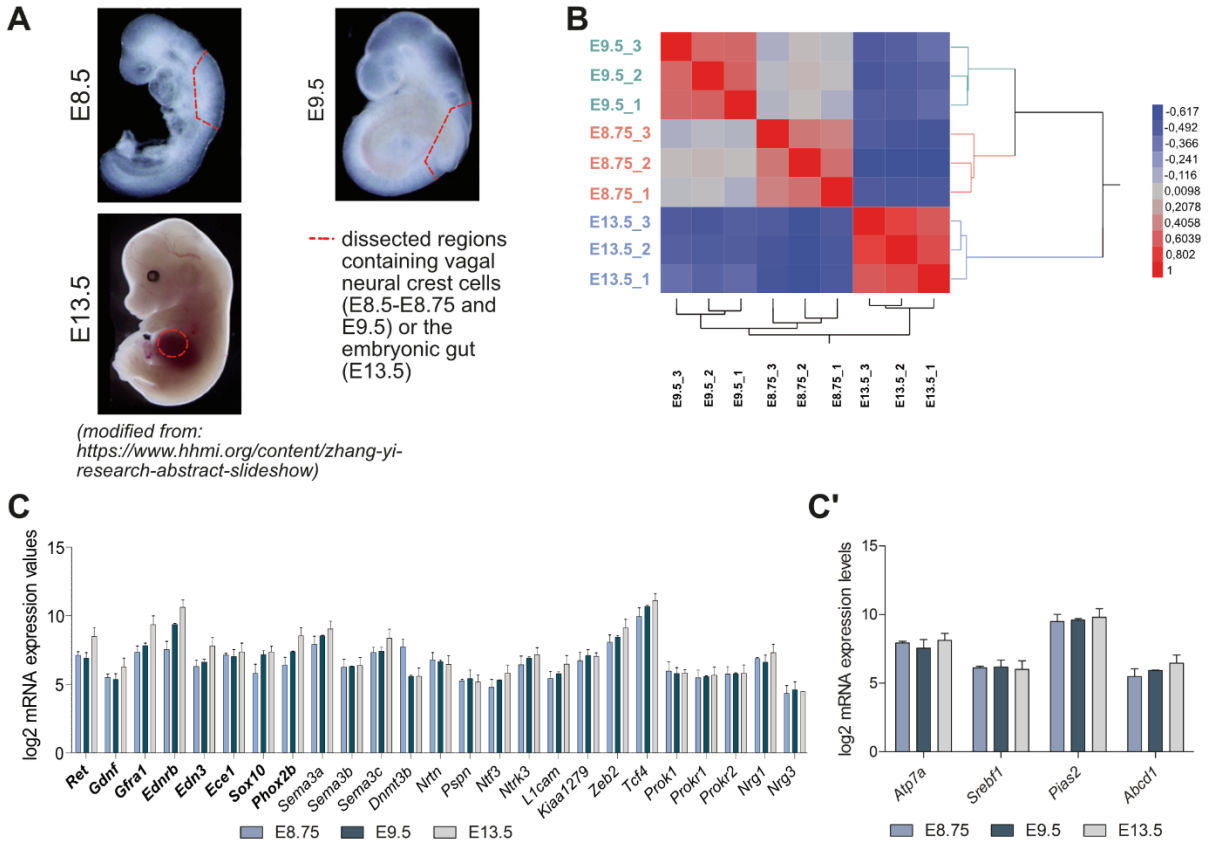


Figure 9: Transcriptomics data generated from murine tissue samples at relevant developmental stages

A: Murine embryonic tissues of three developmental stages (E8.75, E9.5, E13.5) were prepared, whereas pre-migratory vagal NCCs were extracted at stages E8.75 and E9.5, embryonic gut tissue was dissected at stage E13.5. Dissected regions are highlighted in red. Pictures are modified from: https://www.hhmi.org/content/zhang-yi-research-abstract-slideshow. **B:** Transcriptome profiling results of total RNAs extracted from the investigated specimens are illustrated as hierarchical clustered heatmap. Color code represents the Pearson's coefficient (red: high/blue: low similarity between samples). Overall mean expression data of E13.5, E8.75 and E9.5 were used for clustering analysis. **C and C':** Transcriptional profiling of examined embryonic tissues displayed as column bar graphs. Exemplarily, expression profiles of known HSCR risk loci (C) and candidate genes (C') are presented. Most prominent HSCR risk genes are lettered in bold (C). Bar plots show normalized mean + standard error of mean (SEM) (n=3).

showed an increase in mRNA expression during embryonic development from E8.75 to E13.5. An upregulation in the gene expression during development applied as well to the majority of the other HSCR risk genes, while for *Persephin (Pspn)*, all genes of the *Prokineticin (Prok)* family, *Nrg3* and *Sema3b* stable mRNA expression values in all investigated tissues were detectable (fig. 9 C).
The novel candidate genes were expressed at comparable levels as the currently known HSCR risk genes (fig. 9 C‘). Lowest log2 mRNA expression values of 5-6 were detectable for *Srebf1* and *Abcd1* while expression values of *Atp7a* and *Pias2* ranged between 7-10. Each candidate gene was expressed at almost similar levels in the three investigated tissues.

3.2.4 *Expression analyses on protein level*

Murine tissues

Complementary to transcriptional profiling, IF analyses using cryosections of whole mount embryos at corresponding developmental stages (E9.5, E10.5, E11.5, E13.5) were carried out. Immunoreactivity of candidates was investigated in cell lineages relevant for ENS development and function, such as NCCs, neuronal, epithelial as well as smooth muscle cells. In order to assess specific cell types, pre-migratory vagal NCCs at stages E9.5 and E10.5 or ENCDCs at stage E11.5 were visualized by Sox10. Neuronal progeny at stages E11.5 and E13.5 was stained with the neuronal markers Beta III-Tubulin (Tubb3), Ubiquitin C-Terminal Hydrolase L1 (Pgp9.5) or Hu-Antigen C/D (HuC/D). Epithelial as well as smooth muscle cells were identified by Epithelial cadherin (ECad) and Smooth muscle actin (Sma), respectively. Results for all four candidates are described in more detail in the subsequent paragraphs and displayed in the following figures.

Protein expression analyses for **Atp7a** showed an almost ubiquitous expression pattern in all investigated ENS-relevant tissues and stages (fig. 10). An expression of Atp7a and the NCC marker Sox10 within the same cell might be concluded based on antigen specific IF signals in close proximity to each other. This was observable at stages E9.5, E10.5 and E11.5 (fig. 10 A-C, arrows). At E9.5 and E10.5, $Sox10^{+}$ cells located in close proximity to the neck of the embryo were investigated and indicated the presence of pre-migratory vagal NCCs. At higher magnification, signals of the NCC marker and Atp7a were observable in close spatial proximity suggesting antigen co-expression (fig. 10 A and B, arrows). At E11.5, only a small number of invading ENCDCs ($Sox10^{+}$) were detectable but again a partial co-staining with Atp7a might be suggested (fig. 10 C, arrow). Using a set of different neuronal cell fate markers (Tubb3, Pgp9.5, HuC/D), Atp7a was shown to be partially expressed in or in close proximity to these cell types at E11.5 (fig. 10 D’-D’’’, arrows) and E13.5 (fig. 10 E’- E’’’, arrows), respectively. While immature neurons in presented midgut loops were disorganized at E11.5, neuronal cells at E13.5 were already structured in a ring-like shape indicating the premature myenteric plexus. Comparable to this, smooth muscle cells immunoreactive for Sma were arranged in a ring-like structure at E13.5 and showed probably a partial co-expression with Atp7a (fig. 10 D’, arrow). No overlapping expression was observable for Atp7a and the epithelial marker ECad at stages E11.5 (fig. 10 D’’’’) and E13.5 (fig. 10 E’’’’).

The transcription factor **Srebf1** was almost ubiquitously expressed in all investigated tissue regions and stages (fig. 11). Srebf1 was found to be partially co-expressed with the NCC marker Sox10 in vagal pre-migratory NCCs at E9.5 (fig. 11 A, arrow) and E10.5 (fig. 11 B, arrow) as well as in gut invading NCC-derived progenitors at E11.5 (fig. 11 C, arrow). A partial overlap of Srebf1 with the neuronal markers Tubb3, Pgp9.5, HuC/D was detectable at E11.5 (fig. 11 D'-D''', arrows) and E13.5 (fig. 11 E'-E''', arrows), respectively. Additionally, IF analyses showed a co-expression between Srebf1 and the muscular marker Sma at E13.5 (fig. 11 E, E', arrow). While almost all enterocytes within the epithelial lining of the gut lumen ($ECad^+$) at stage E11.5 co-expressed Srebf1 (fig. 11 D'''', arrow), this ubiquitous expression pattern of Srebf1 in the epithelial layer seemed to be more restricted at E13.5 (fig. 11 E'''', arrow). However, a partial overlap in the expression of ECad and Srebf1 was still detectable for epithelial cells at the basal side.

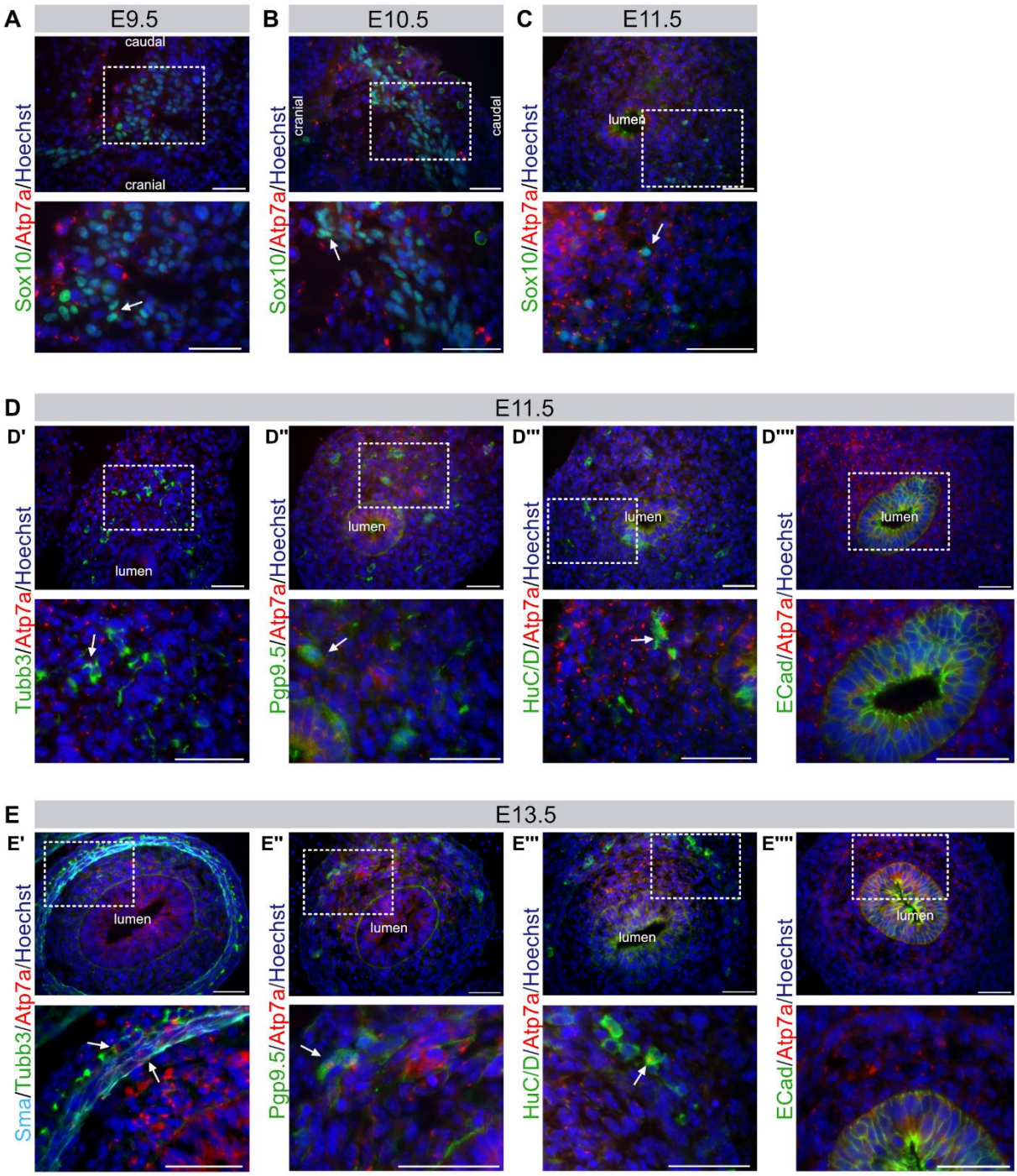


Figure 10: Expression analyses of Atp7a in cryosectioned whole mount embryos

A-C: Regions harbouring pre-migratory vagal NCCs in close proximity to the neck of the embryo, at stages E9.5 (A) and E10.5 (B), and a midgut loop at stage E11.5 (C) were stained for the candidate Atp7a (in red) and the NCC marker Sox10 (in green). Cranial and caudal positions (A, B) and the gut lumen are indicated (C). IF signals in close spatial proximity to each other suggesting a co-expression of the antigens are highlighted by white arrows. **D:** Expression analyses of Atp7a (in red) in the developing gut at stage E11.5 using several neuronal markers (Tubb3 (D'), Pgp9.5 (D''), HuC/D (D''')) as well as the epithelial marker ECad (D'''') (all in green) for co-labelling. Midgut loops with gut lumen are displayed. Atp7a expressing cells presumably of neuronal origin are highlighted by white arrows while no epithelial cells display immunoreactivity for the candidate. **E:** At stage E13.5, co-labelling with Atp7a (in red) and cell type specific markers (Tubb3 (in green) and Sma (in cyan) (E'); Pgp9.5 (E''), HuC/D (E'''), ECad (E'''') (all in green)) were performed. Midgut loops with lumen are shown. Double positive cells are highlighted by white arrows but were only detectable for neuronal and muscular lineage.
Lower images show magnifications of the areas highlighted in the upper images. Nuclei were counterstained with Hoechst (blue). Scale bar: 50µm.

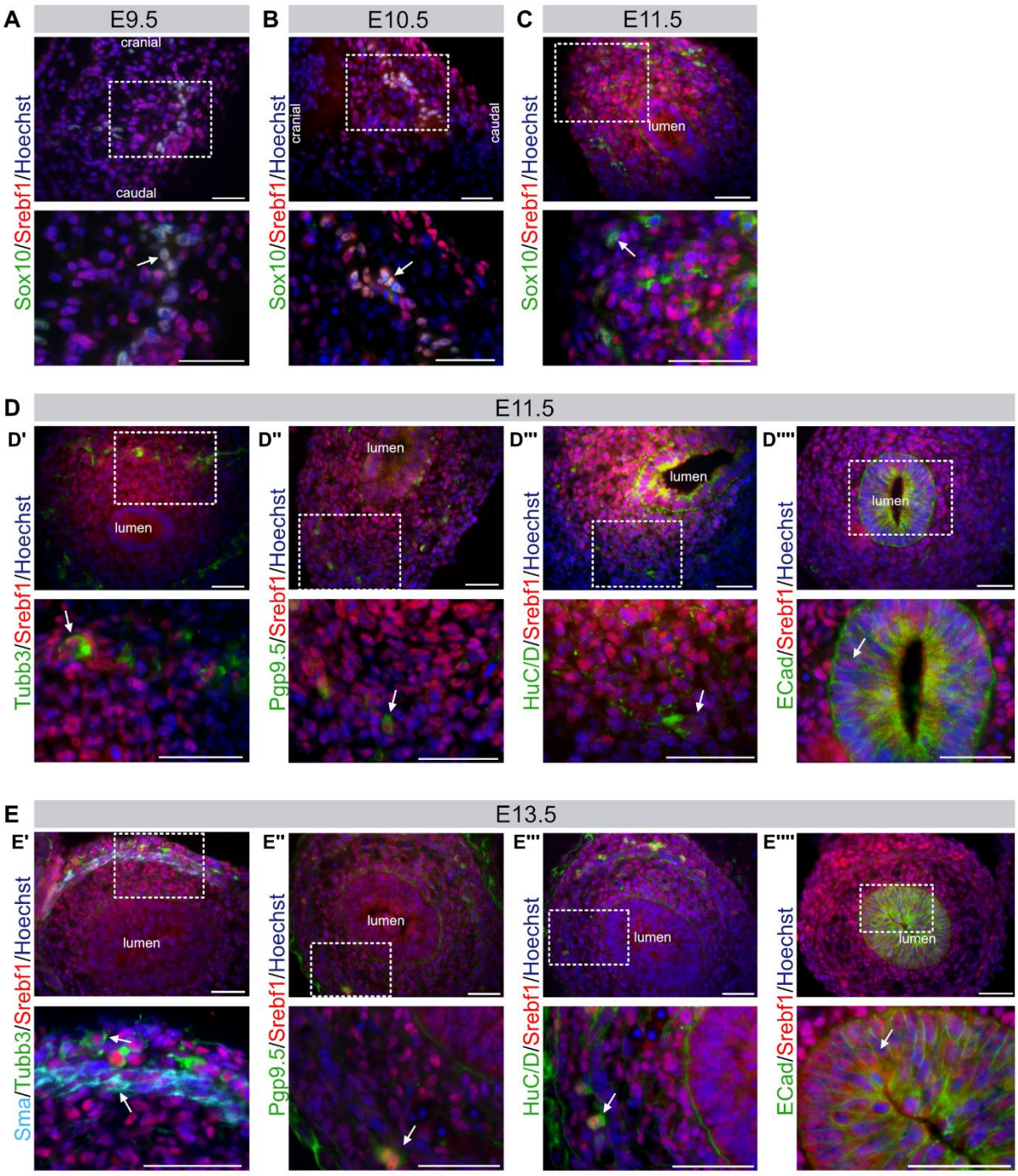


Figure 11: Immunocytochemical analyses of Srebf1 in whole mount embryo sections

A-C: Double labelling of Srebf1 (in red) and Sox10 (in green) at stage E9.5 (A), E10.5 (B) or E11.5 (C). Co-expressing cells are marked by white arrows. Pre-migratory vagal NCCs ($Sox10^{+}$) located in close proximity to the neck of the embryo were analysed; cranial and caudal positions are indicated (A, B). A midgut loop with gut lumen is shown for E11.5 (C). **D:** Co-staining of Srebf1 (in red) with various cell type specific markers (neuronal: Tubb3 (D'), Pgp9.5 (D''), HuC/D (D'''), epithelial: Ecad (D'''')) (all in green) at stage E11.5. Midgut loops with gut lumen are displayed. Immature neurons and epithelial cells show co-expression with Srebf1 as highlighted by white arrows. **E:** Visualization of Srebf1 (in red) expression in the midgut at E13.5 (lumen is labelled). Triple (Tubb3 (in green), Sma (E') (in cyan)) or double-staining (Pgp9.5 (E''), HuC/D (E'''), ECad (E'''') (all in green)) are shown. White arrows highlight co-stained cells.

Lower images are magnifications of highlighted regions in the upper images. Nuclei were counterstained with Hoechst (blue). Scale bar: 50µm.

Abcd1 was shown to be almost ubiquitously expressed in relevant ENS tissues of early stages (fig. 12 A and B) but its expression was slightly more restricted to individual cells at later phases of GI tract formation (fig. 12 C-E). In general, the staining pattern suggested a Golgi or vesicular specific protein localisation. IF signals of Abcd1 and the NCC marker Sox10 resided in close spatial proximity to each other at E9.5 (fig. 12 A, arrow), at E10.5 (fig. 12 B, arrow) and at E11.5 (fig. 12 C, arrow), and therefore pointed towards a co-expression of both proteins at these stages. Staining of the neuronal progeny at E11.5 using the neuronal markers Tubb3, Pgp9.5 and HuC/D (fig. 12 D'-D''', arrows) in the developing midgut led to comparable observations as previously described. A positive immunoreactivity of Abcd1 was also observable at E11.5 in ECad$^+$ epithelial cells of the developing gut mucosa (fig. 12 D'''', arrow). At stage E13.5, a partial co-expression of Abcd1 in immature neuronal cells (Tubb3$^+$, Pgp9.5$^+$, HuC/D$^+$) (fig. 12 E'-E''', arrows), epithelial cells (ECad$^+$) (fig. 12 E'''', arrow) as well as smooth muscle cells (Sma$^+$) (fig. 12 E', arrow) might be concluded based on a partial overlap in respective antigen specific IF signals.

The expression pattern of **Pias2** pointed towards a cytoplasmic protein localisation (fig. 13 A-E). Comparable to Srebf1, Pias2 was also almost ubiquitously expressed in all investigated tissues of interest and developmental stages. Sox10$^+$ vagal NCCs located in close proximity to the otic vesicle at E9.5 and E10.5 displayed Pias2-specific IF signals in their spatial proximity (fig. 13 A and B, arrows). This might indicate a co-expression of both proteins. Comparable observations were documented for gut invading ENCDCs (Sox10$^+$) at E11.5 (fig. 13 C, arrow) suggesting that these cells maintain candidate protein expression. IF analyses showed Pias2 immunoreactivity in close proximity to immature neurons of the developing gut at E11.5 labelled with Tubb3, Pgp9.5 and HuC/D (fig. 13 D'-D''', arrows). Additionally, an epithelial cell specific expression for Pias2 was observable at this stage (fig 13 D'''', arrow) and at E13.5 (fig.13 E'''', arrow) by co-staining with ECad. At E13.5, again a putative co-staining of Pias2 with the neuronal cell type markers (Tubb3, Pgp9.5, HuC/D) (fig. E'-E''', arrows) and the smooth muscle marker (Sma) (fig. 13 E', arrow) might be suggested based on a putative, partial overlap of the respective IF signals.

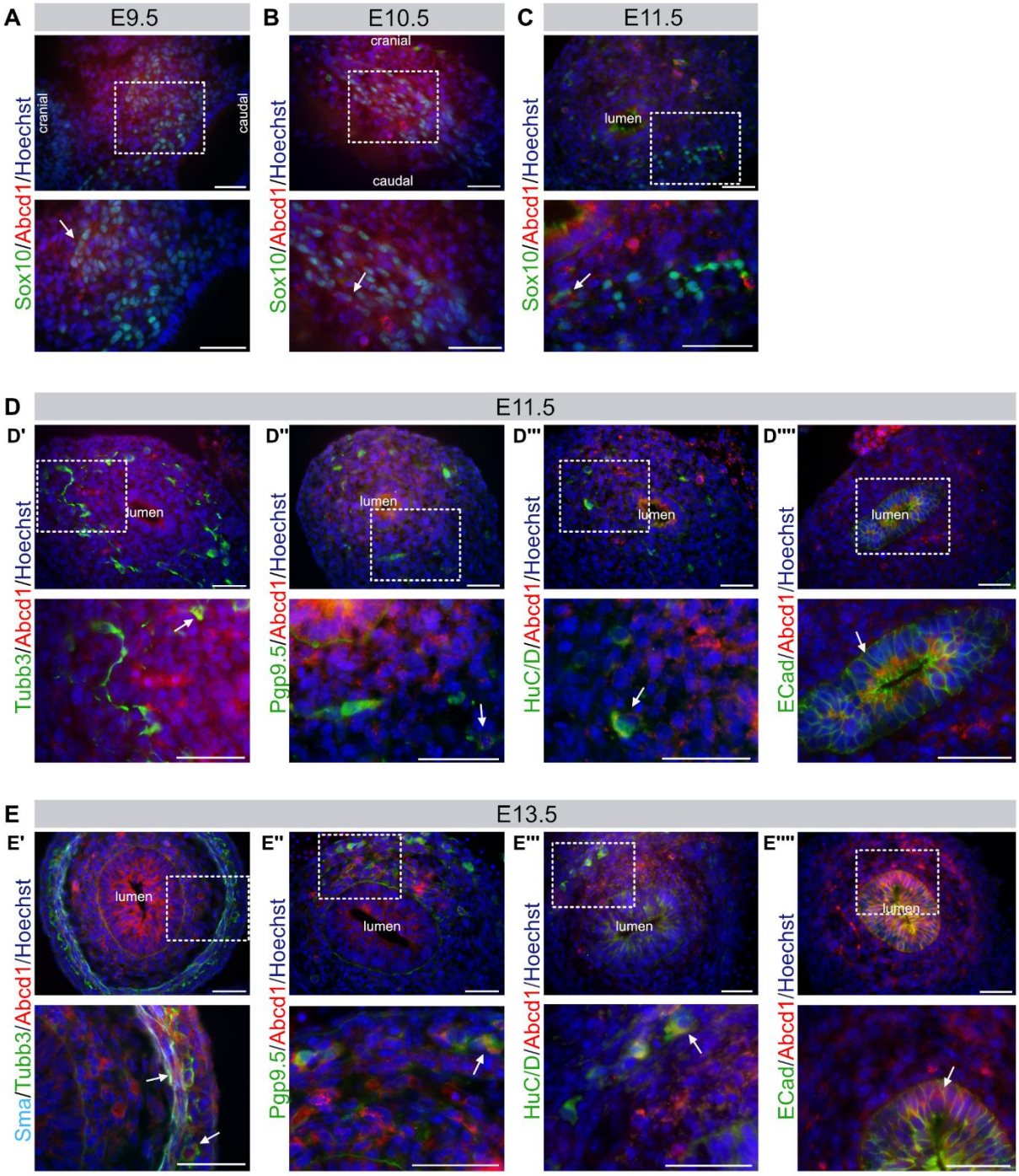


Figure 12: Expression analyses of Abcd1 in murine embryos

A-C: Immunocytochemical analysis of Abcd1 (in red) and Sox10 (in green) in cryosectioned embryos of stage E9.5 (A), E10.5 (B) and E11.5 (C). At E9.5 and E10.5, $Sox10^{+}$ cells caudally to the otic vesicle are shown and cranial and caudal positions are illustrated (A, B). At E11.5, a midgut loop with gut lumen is displayed (C). Arrows highlight regions where IF signals are detectable in close spatial proximity to each other indicating a putative co-expression of both antigens. **D:** Expression analyses of Abcd1 (in red) in the developing gut at stage E11.5 using different cell type specific markers for co-labelling (immature neurons are stained with Tubb3 (D'), Pgp9.5 (D''), HuC/D (D'''), epithelial cells are stained with ECad (D'''')) (all in green). Midgut loops with gut lumen are displayed. Individual cells or regions showing immunoreactivity for Abcd1 and the cell type specific marker are highlighted by an arrow. **E:** IF analyses of Abcd1 (in red) and a set of cell type specific markers (Tubb3 (in green), Sma (E') (in cyan); Pgp9.5 (E''), HuC/D (E'''), ECad (E'''') (all in green)) in the developing gut at stage E13.5. The gut lumen is indicated. Putatively Abcd1 expressing neuronal, muscular or epithelial cells as concluded by means of partially overlapping IF signals are marked (white arrow).

Magnifications of selected regions are displayed in the lower images. Nuclei were counterstained with Hoechst (blue). Scale bar: 50µm.

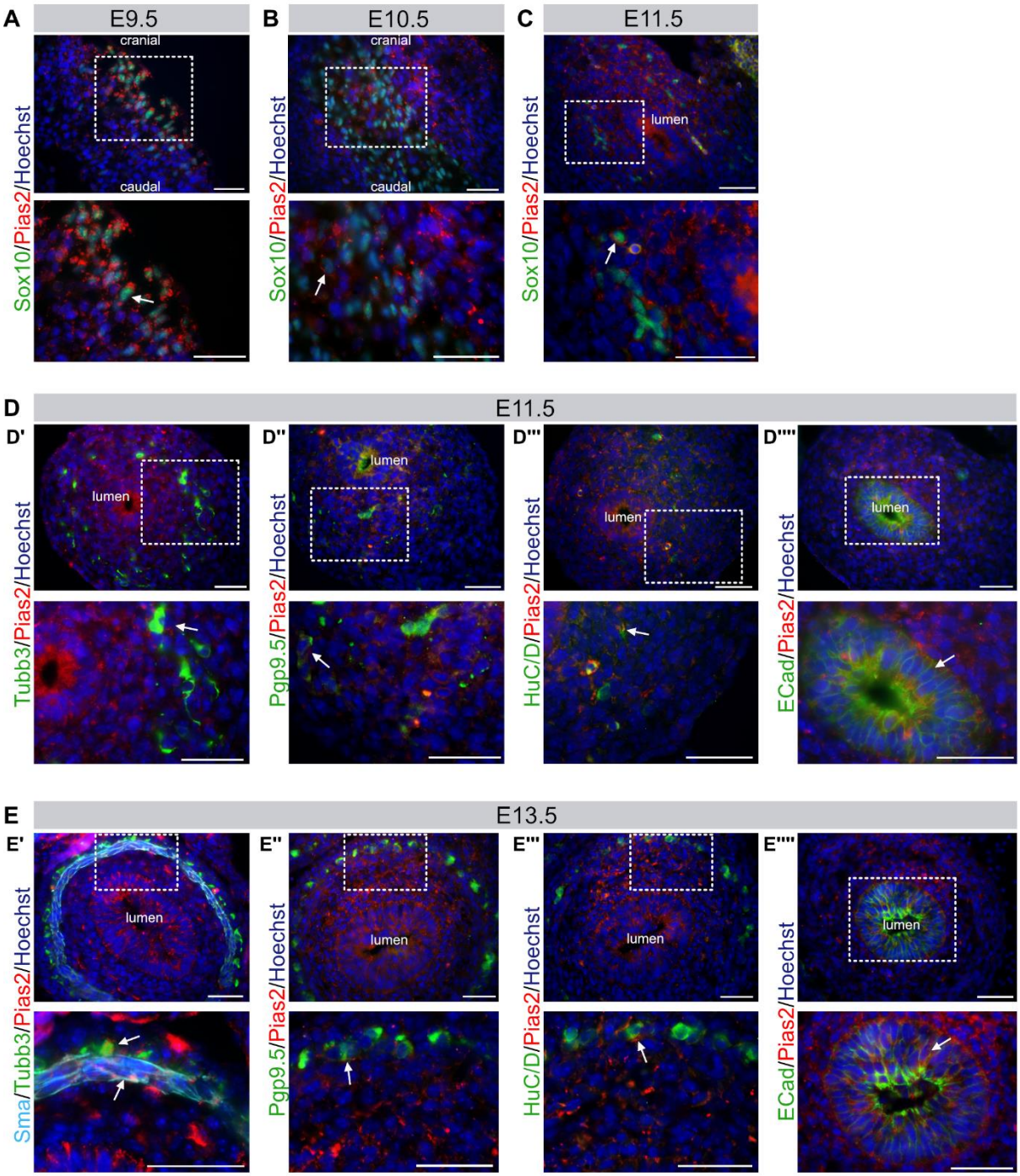


Figure 13: IF staining of Pias2 in cryosectioned murine embryos at several developmental stages

A-C: IF analyses of Pias2 (in red) and Sox10 (in green) at early developmental stages E9.5 (A), E10.5 (B) and E11.5 (C). While pre-migratory vagal NCCs at the respective position are shown for E9.5 and E10.5 (A, B), $Sox10^+$ cells in the developing midgut at E11.5 are displayed (C). Cranial and caudal positions (A, B) and the gut lumen (C) are highlighted. IF signals for the candidate and Sox10 reside in close spatial proximity to each other as indicated by white arrows. **D:** At stage E11.5 expression analyses of Pias2 (in red) with respective neuronal markers (Tubb3 (D'), Pgp9.5 (D''), HuC/D (D''')) and the epithelial marker ECad (D'''') (all in green) were performed. Midgut loops with gut lumen are displayed. Cells or regions immunoreactive for both investigated antigens are highlighted (white arrow). **E:** Pias2 (in red) expression in the developing gut at stage E13.5. The gut lumen is indicated. Co-labelling of the candidate was performed with different cell type specific markers (Tubb3 (in green), Sma (E') (in cyan); Pgp9.5 (E''), HuC/D (E'''), ECad (E'''') (all in green)). Presumably, Pias2 co-expressing cell types concluded based on partially overlapping IF signals, are marked (white arrow).

Lower images show magnifications of the regions highlighted in the upper images. Nuclei were counterstained with Hoechst (blue). Scale bar: 50µm.

Complementary immunocytochemical analyses were performed in adult murine colon cryosections assessing the expression of candidates and respective cell type specific markers in enteric ganglia. All candidates were shown to be expressed in neuronal cells of the ganglionated plexus by co-staining with Tubb3 or HuC/D, respectively (fig. 14). IF analyses of Atp7a in adult murine colon sections revealed prominent IF signals in the cell soma of Tubb3$^+$ or HuC/D$^+$ neurons (fig 14 A and A'). A nuclear restricted staining was detectable for Srebf1. Srebf1$^+$ cells partially co-expressed the neuronal markers Tubb3 and HuC/D. In addition, Srebf1 signals were as well detectable in cells not being immunoreactive for the depicted neuronal markers but residing in close proximity to enteric ganglia (fig. 14 B and B'). A partial overlap in the expression of Tubb3 or HuC/D with the candidates Abcd1 (fig. 14 C and C') and Pias2 (fig. 14 D and D') was as well demonstrated in the respective analyses. Both candidates indicated a dotted-like staining pattern.

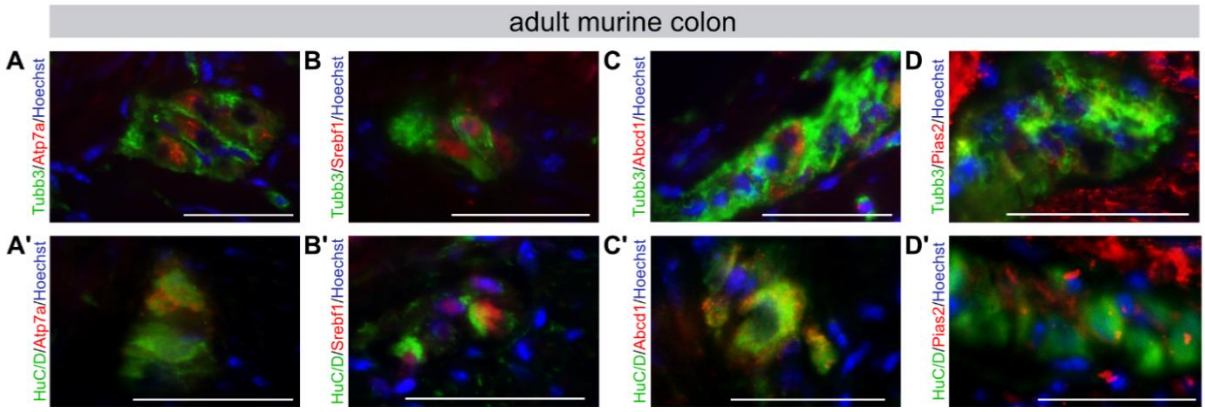


Figure 14: Expression analyses of candidates in adult murine colon sections

A-D': IF analyses of candidates (all in red) co-stained with the neuronal markers Tubb3 (A-D) and HuC/D (A'-D') (both in green) in adult murine colon cryosections (Atp7a (A, A'), Srebf1 (B, B'), Abcd1 (C, C'), Pias2 (D, D')). Nuclei were counterstained with Hoechst (blue). Scale bar: 50µm.

To assess candidate specific immunoreactivity as well as expression levels in different murine embryonic and adult GI and brain tissues, Western blot analyses were carried out (fig. 15 A-D). Protein lysates of HEK293TN cells, transiently transfected with candidate gene-specific overexpression constructs, were used as internal controls.

Western blot analyses for Atp7a showed multiple bands between 35-50kDa. The Atp7a-specific signal at around 190kDa was confirmed based on protein lysate of the positive control (HEK293TN transiently transfected with *ATP7A* MYC-tagged overexpression construct). In general, highest protein expression levels for Atp7a (predicted size: 163kDa) were observable in embryonic gut and brain tissues. Minimal protein expression based on the presence of faint bands at the respective size could be concluded for adult duodenum and adult colon specimens. In adult jejunum/ileum, brain and cortex samples, an expression of Atp7a was not detectable (fig. 15 A).

For Srebf1, again embryonic tissues showed highest protein expression levels for the full-length (FL) (predicted size: 121kDa) and the mature (MAT) protein variant (predicted size: 70kDa). All adult specimens containing ENS or CNS relevant cell types displayed only marginal expression of the MAT Srebf1 protein form but no expression of the FL variant. Srebf1-specific signals in the samples of interest were confirmed based on protein lysates of the two positive controls as in general a multiple of bands

was observable in Western blot analyses. While the transient transfection of the FL overexpression construct led to the generation of one Srebf1-specific band in the respective protein lysate, the control sample transfected with the MAT overexpression construct displayed three signals of variable sizes ranging between 70-110kDa (fig. 15 B).

For Abcd1, GI tissues showed most prominent protein expression in general. Specifically, expression was detected to be highest in adult duodenum and lowest in adult jejunum/ileum. In addition to a band at around 80kDa indicating for the monomeric form of Abcd1, a signal at around 130kDa was observable in all adult GI tissues, which most probably represent the dimeric protein form. Due to the Green fluorescent protein (GFP) tag, recombinantly expressed ABCD1 (in HEK293TN cells) indicated a larger protein size than endogenously expressed murine Abcd1. Expression of Abcd1 in adult brain tissues and the embryonic brain specimen was weak but detectable except for the adult cortex specimen (fig. 15 C).

Western blot analyses for Pias2 indicated an endogenous expression of the two major isoforms Pias2ß (predicted size: 68kDa) and Pias2α (predicted size: 63kDa) with highest levels in embryonic gut and brain specimens. Besides these Pias2-specific signals at around 70-80kDa as validated by means of the positive control protein lysate, multiple other bands of lower sizes were observable in immunoblot analyses. Except for adult duodenum, where no endogenous candidate protein expression was detectable, all other investigated tissues displayed expression of at least one isoform but at low levels. The α isoform was found to be predominantly expressed in brain tissues as adult brain and adult cortex did not display an endogenous expression of the ß isoform (fig. 15 D).

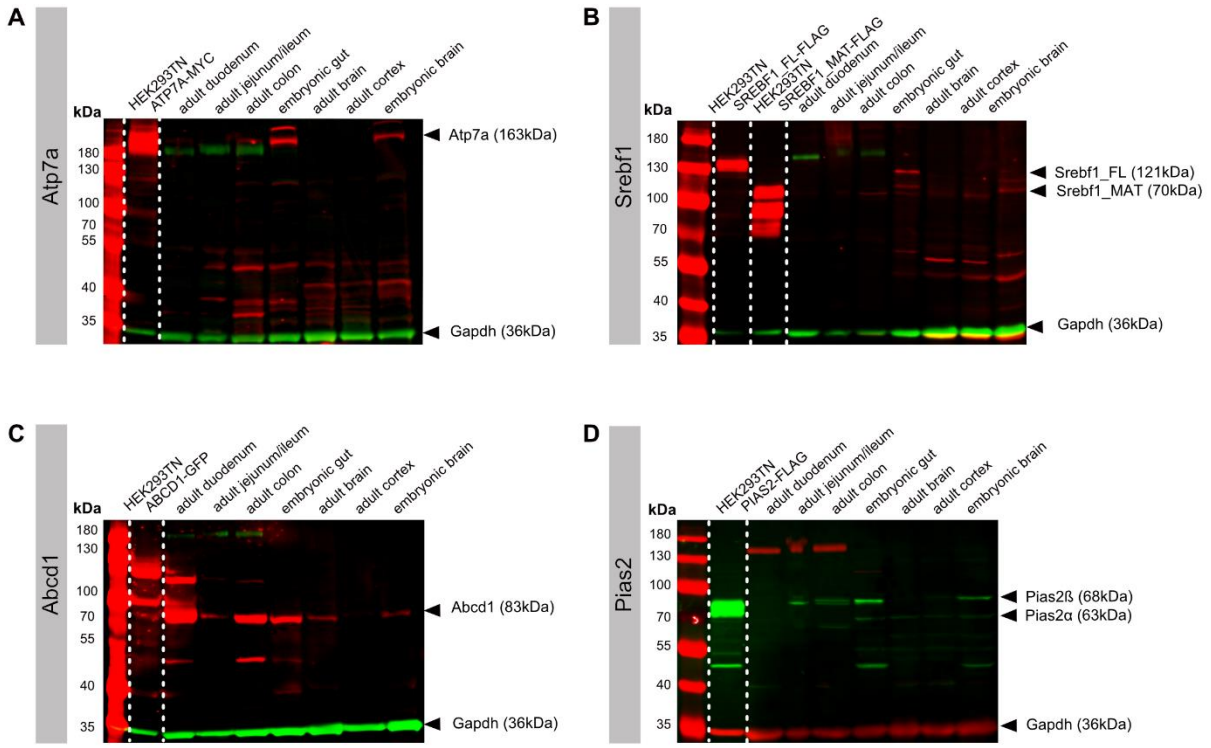


Figure 15: Western blot analyses in murine GI and brain tissues

A-D: Candidate specific protein expression was assessed in different GI (adult: duodenum, jejunum/ileum, colon; embryonic (E15.5): gut) and brain tissues (adult: brain, cortex only; embryonic (E15.5): brain) (Atp7a (A), Srebf1 (B), Abcd1 (C), Pias2 (D)). Protein lysates of HEK293TN cells transiently transfected with respective candidate specific overexpression constructs were implemented as antibody positive controls. Gapdh was used as loading control. Predicted protein sizes are annotated. FL (full-length), MAT (mature)

Human tissues

Complementary IHC staining of all four candidates was performed using FFPE sections of human fetal and adult colon.

In the fetal colon, ATP7A and PIAS2 were shown to be prominently expressed within enteric ganglia (fig. 16 A, D). In contrast, SREBF1 was only marginally expressed in enteric ganglia and surrounding tissue displayed comparable signal intensities (fig. 16 B). ABCD1 was not expressed in respective neuronal tissue structures (fig. 16 C). However, all candidates showed a positive immunoreactivity in the gut epithelium. ATP7A displayed an almost ubiquitous expression in cells of the epithelial lining and the lamina propria while SREBF1 was distinctly expressed in individual cells at the crypt's base (fig. 16 A' and B'). Comparable to ATP7A, ABCD1 showed as well an almost ubiquitous expression presumably in immune cells within the lamina propria (fig. 16. C'). For PIAS2, immunocytochemical signals were observable in the cells located at the base of the crypts (fig. 16 D').

In the adult colon, protein expression patterns of ATP7A and PIAS2 were highly comparable to those observed in fetal tissues as both candidates were strongly expressed in enteric ganglia as well as in epithelial cells of the crypts and cells of the lamina propria (fig. 16 E, E', H, H'). In contrast to the fetal stage, SREBF1 showed a distinct staining pattern in individual cells of the adult enteric ganglia. Moreover, prominent signals were as well observable in individual cells of the lamina propria (fig. 16 F and F'). This was comparable to observations for ABCD1 as specific signals were as well detectable

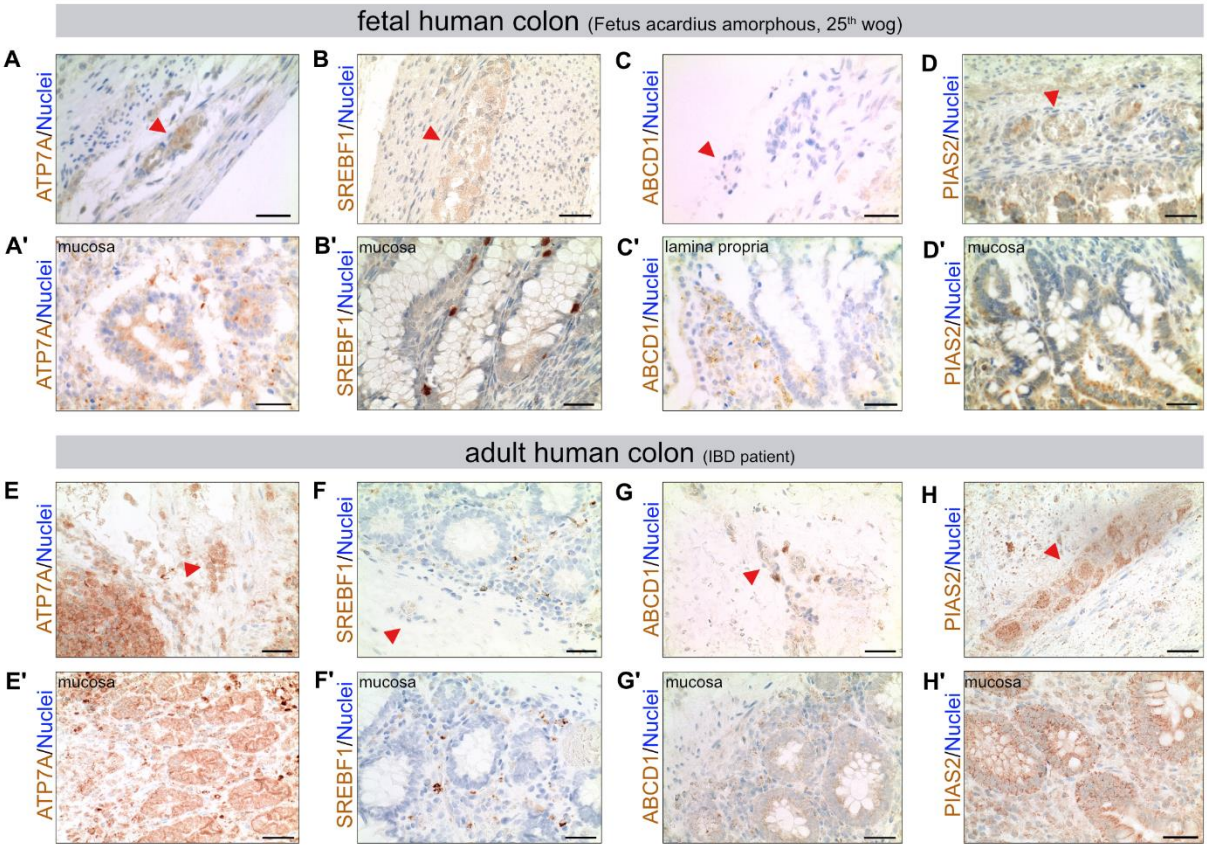


Figure 16: IHC staining of candidates in fetal and adult human colon sections

A-D: IHC staining of candidates (ATP7A (A), SREBF1 (B), ABCD1 (C), PIAS2 (D)) in human fetal colon tissue originating from a fetus acardius amorphous. **E-H:** IHC staining of candidates (ATP7A (E), SREBF1 (F), ABCD1 (G), PIAS2 (H)) in human adult colon tissue of an IBD patient.

Upper images show enteric ganglia highlighted by a red triangle (A-H). Lower images show epithelial tissue layers (A'-H'). Candidates (DAB brown staining). Nuclei were counterstained with Hematoxylin (blue). Scale bar: 20µm.

within adult enteric ganglia and cells of the lamina propria (fig. 16 G, G'). In contrast to the SREBF1-specific staining, ABCD1 expression in the epithelium was more ubiquitous as cells at the crypt's base showed an immunoreactivity for this candidate too.

3.3 Candidate gene characterization

For candidate gene characterization on functional level, the generation of gene-specific knockout (*KO*) cell clones applying the CRISPR/Cas9 technology to the human neuroblastoma cell line SHSY5Y was carried out. Except for the four candidate genes, the major HSCR susceptibility locus *RET* was implemented as *proof-of-principle* control into the study pipeline. A control transfected and Fluorescent-activated cell (FAC) sorted *mock control* clone was already generated in a previous study and served as internal reference (Master Thesis Mederer, 2015).

3.3.1 *Genome editing*

Endogenous expression of all targeted genes within the chosen cell source was demonstrated prior to editing (fig. S4, Supplementary). FAC sorted single cell clones were evaluated according to their cell morphology by brightfield microscopy as only clones with neuronal-like morphology were included for genotyping and further investigations.

The ***RET***-specific sgRNA targeted exon 2, whereas Cas9 mediated DNA double strand break caused the homozygous insertion of a cytosin at the respective position of the protospacer adjacent motif (PAM) (c.231_232insC) (fig. 17 A'; fig. S5, Supplementary) (Master Thesis Mederer, 2015). The CRISPR/Cas9 induced KO of RET was validated by Western blot analysis as no RET protein was detectable at the corresponding size (predicted size: 124kDa) in the clone (p.T78Hfs5*). Furthermore, no shortened protein version of RET was observable. A FLAG-tagged *RET* overexpression construct transiently transfected in HEK293TN cells was used as control in immunoblot analyses. Endogenous expression levels of RET in the parental cell line SHSY5Y and the *mock control* clone were comparable (fig. 17 A'').

The sgRNA targeting ***ATP7A*** was specific to exon 3. The generated *ATP7A KO* clone carried a compound heterozygous genome modification at the probable Cas9 cutting site as validated by Sanger sequencing (c. 500_501insT/500_501insA) (fig. 17 B'; fig. S5, Supplementary). In Western blot analyses, total protein lysates of HEK293TN cells transiently transfected with a MYC-tagged *ATP7A* overexpression construct, lysates of SHSY5Y cells and of the *mock control* clone served as controls. All investigated samples except for the *ATP7A KO* clone (p.Q167Hfs*4/p.A168Sfs*3) showed a band at 190kDa indicating ATP7A was successfully knocked-out in the genome edited clone. A truncated protein was not detectable in total protein lysate of the *ATP7A KO* clone either. Endogenous ATP7A expression in the *mock control* clone was higher than in the parental cell line SHSY5Y (fig. 17 B'').

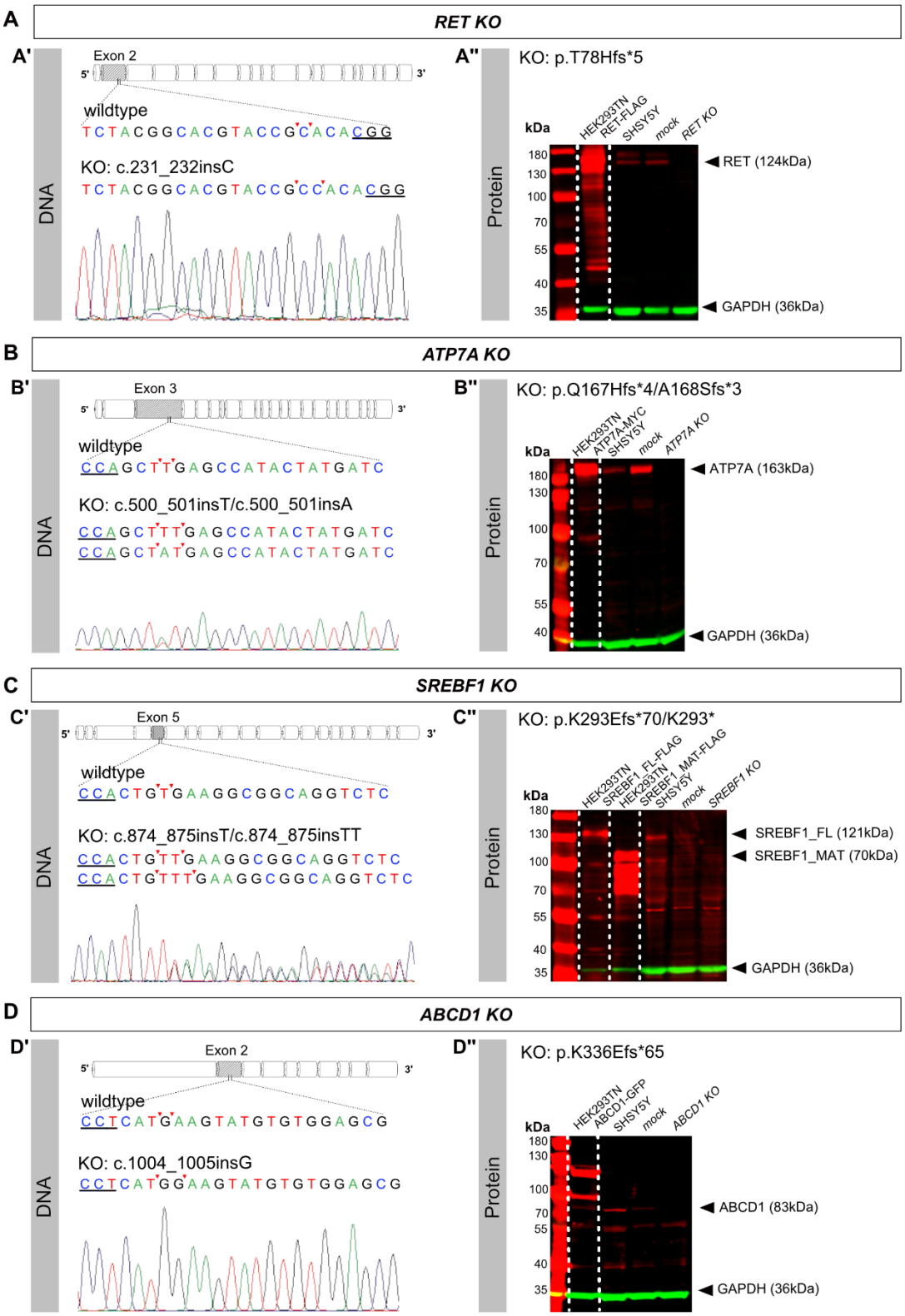


Figure 17: CRISPR/Cas9 gene mediated knockout of RET and candidate genes in the neuroblastoma cell line SHSY5Y

A'-D': Gene-specific sgRNAs were designed against marked exons (in grey). Cas9 mediated double-strand breaks were repaired by non-homologous end joining (NHEJ) causing homozygous or compound heterozygous genome modifications at the respective positions as verified by Sanger sequencing. Four coloured sequencing chromatograms are shown for each gene (*RET* (A'), *ATP7A* (B'), *SREBF1* (C'), *ABCD1* (D')). PAM sites are underscored. Putative Cas9 cutting sites are highlighted by red arrowheads. **A''-D'':** Knockout (KO) on protein level was validated by Western blot analyses using different protein lysates as internal controls (HEK293TN cells transiently transfected with a gene-specific, tagged overexpression construct, SHSY5Y cells, *mock control* cells) (RET (A''), ATP7A (B''), SREBF1 (C''), ABCD1 (D'')). GAPDH was used as loading control in all Western blot analyses. Predicted protein sizes are annotated. Images show modified blots as individual lanes of one blot were rearranged if necessary. FL (full-length), MAT (mature)

Genome editing of ***SREBF1*** was achieved by using a sgRNA specific to exon 5. Allele specific cloning and respective Sanger sequencing confirmed a compound heterozygous genome modification (c.874_875insT/c.874_875insTT), introduced 3-4 nucleotides upstream of the PAM (fig. 17 C'; fig. S5, Supplementary). The CRISPR/Cas9 induced SREBF1 KO could be also confirmed by Western blot analysis as no bands indicating FL or MAT SREBF1 were observable for the respective *KO* clone (p.K293Efs*70/K293*). Protein lysates from HEK293TN cells transiently expressing the FL and the MAT form of SREBF1 as well as from SHSY5Y cells and the *mock control* clone were used as positive controls. In general, a multitude of bands was detectable in immunoblot analyses. Endogenous expression levels of the FL SREBF1 (predicted size: 121kDa) in the *mock control* clone was lower than in SHSY5Y cells. A band indicating for the MAT form of SREBF1 (predicted size: 70kDa) was not detectable in the total protein lysate of the *mock control* clone (fig. 17 C'').

The candidate gene ***ABCD1*** was CRISPR/Cas9 edited in exon 2 via the homozygous insertion of a guanosin nucleotide at the respective DNA cutting site (c.1004_1005insG) (fig. 17 D'; fig. S5, Supplementary). On protein level, Western blot analyses demonstrated the complete protein loss as no specific band at 83kDa or any additional band at a lower protein size was observable in total protein lysate of the *ABCD1 KO* clone (p.K336Efs*65). Protein lysates of HEK293TN cells transiently transfected with a GFP-tagged *ABCD1* overexpression construct as well as of SHSY5Y cells and the *mock control* clone served as controls (fig. 17 D'').

Results for forward and reverse Sanger sequencing of the edited genomic region for all clones are shown in the supplementary data (fig. S5). In addition, all *KO* clones were evaluated but negatively tested for alternative splicing events of the genome edited exon on transcript level (fig. S6, Supplementary). Additionally, selected off-target regions were assessed for unintended genomic modifications. Except for a heterozygous one-nucleotide insertion in the intronic region of *EUKARYOTIC TRANSLATION ELONGATION FACTOR 1 ALPHA 2* (*EEF1A2)* in the *SREBF1 KO* clone, no off-target effects were identified in any of the other clones. As *in silico* investigations did not suggest any functional relevance of this heterozygous variant in the *EEF1A2* intron, the *SREBF1 KO* was validated for further utilization (fig. S7, Supplementary).

For genome editing of *PIAS2*, different sgRNAs were applied in individual experiments (fig. 18). Targeting of exon 6 in a first setup, led to the homozygous insertion of an adenosin nucleotide at the putative Cas9 cutting sites (c.774_775insA) (Master Thesis Mederer, 2015). Alternative splicing of exon 6 was demonstrated by conventional reverse transcription PCR on transcript level using exon 6 flanking primers. Specifically, primers resided in the adjacent exons 5 and 7. Gelelectrophoresis analysis of the generated PCR products demonstrated that only for WT SHSY5Y cells an amplicon of 319bp containing exon 6 was generated. In contrast, for the edited clone a PCR product of 184bp indicating a *PIAS2* amplicon with skipped exon 6 was detectable (fig. 18 A). Sequence identity of the alternatively spliced transcript in the *PIAS2* edited clone was confirmed by Sanger sequencing.

In a second experimental approach, sgRNAs for *PIAS2* were designed specific to exon 2 (fig. 18 B). In total, three different sgRNAs (I-III) were used in individual experiments. After single cell cloning, only clones of epithelial-like morphology survived if sgRNA I was used. These clones were excluded from genotyping and further analyses. Transfection of sgRNAs II and III, resulted as well almost exclusively

in epithelial-like cell clones after FACS. Editing with sgRNA II led to the generation of a clone with neuronal-like morphology harbouring an inserted fragment of the CRISPR/Cas9 cloning vector backbone in its genome and was therefore excluded. A putatively edited clone of neuronal-like morphology was generated after transfection with sgRNA III. Genotyping was not possible as no PCR product spanning the putatively edited region could be generated. Further expression analyses on transcript level revealed the complete absence of the longest annotated isoform (*PIAS2ß*, 1866 bp) in this putatively edited clone.

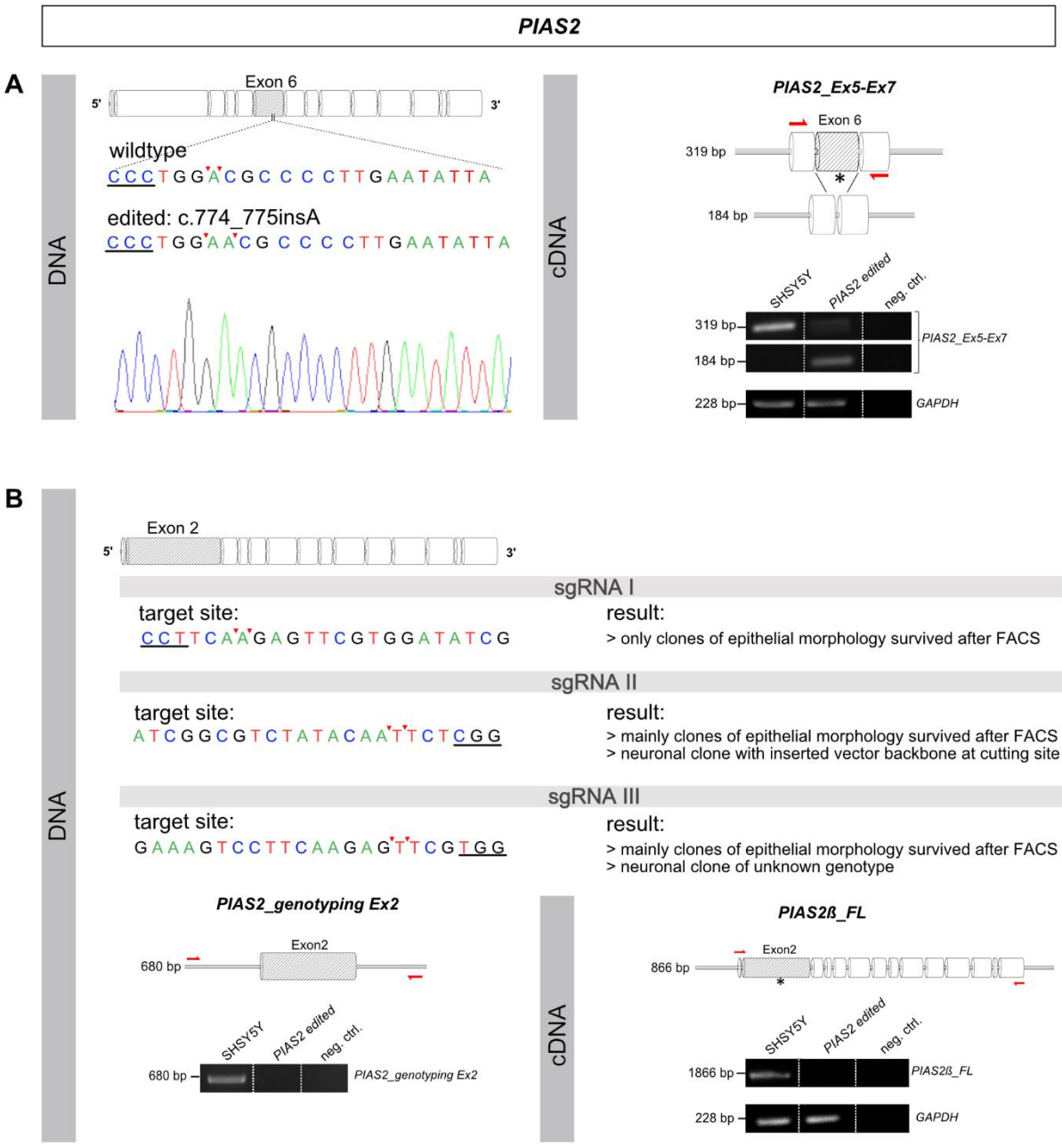


Figure 18: Genome editing approaches for PIAS2

A: sgRNAs specific to exon 6 were used in a first approach for genome editing of *PIAS2* (Master Thesis Mederer, 2015). Four coloured sequencing chromatogram of the edited clone is shown. Analyses of alternative splicing were carried out by conventional PCR on CDAN level. This analysis shows skipping of exon 6 in the edited *PIAS2* clone, as a PCR product of 184bp but not of 319bp (as for WT SHSY5Y cells) was detectable. **B:** Different sgRNAs specific to exon 2 (I-III) were used for genome editing of *PIAS2* in a second approach. Conventional reverse transcription PCR was carried out for transcript detection and shows the complete absence of the *FL PIAS2ß* transcript in the clone edited with sgRNA III. PAM sites are underscored. Putative Cas9 cutting sites are highlighted by red arrowheads. Asterisks indicate putatively CRISPR/Cas9 introduced genome modifications. Primers are illustrated by red arrows. GAPDH was used as a control gene for RNA integrity. Expected and observed amplicon sizes are annotated. FL (Full-length)

3.3.2 *Establishment of a neuronal differentiation protocol*

As outlined in section 2.2.2, the establishment of a neuronal differentiation protocol for SHSY5Y cells was based on results generated by two pilot experiments (A, B) (fig. 3). Differentiating cells were monitored by brightfield microscopy and assessed for their expression profile by qRT PCR using a set of various cell type markers.

Pilot experiment A

In general, no expression of the stem/progenitor marker *SOX10* was detected at any investigated time point of differentiation using any of the differentiation protocols (Ia, IIa, IIIa) (fig. 19).

WT SHSY5Y cells, differentiated according to **protocol Ia**, displayed lower expression levels for the neuronal progenitor markers *LOW-AFFINITY NEUTROPHIN RECEPTOR P75NTR (P75NTR)* and *NESTIN (NES)* with advancing differentiation. The neuronal cell fate marker *TACHYKININ PRECURSOR 1 (TAC1)* was not expressed throughout differentiation while the *MICROTUBULE ASSOCIATED PROTEIN TAU (TAU)* showed only minimal expression levels at early differentiation time points. The pan-neuronal markers *UBIQUITIN C-TERMINAL HYDROLASE L1 (UCHL1)* and *TUBB3* were almost stably expressed in all investigated samples. In contrast, expression levels of *ACHAETE-SCUTE FAMILY BHLH TRANSCRIPTION FACTOR 1* (*ASCL1)* first increased after 7+3d DM II and then decreased again with further differentiation. During neuronal cell fate induction, expression levels for the more advanced neuronal cell fate markers *MICROTUBULE-ASSOCIATED PROTEIN 2 (MAP2)*, *SYNAPTOPHYSIN (SYP)*, *GROWTH-ASSOCIATED PROTEIN 43 (GAP43)* and *NPY* gradually increased. Compared to all assessed markers, *NEUROPEPTIDE Y (NPY)* showed the highest expression values with peak expression at differentiation stage 7+14d DM II. In contrast, *MAP2*, *SYP* and *GAP43* were only expressed at low to intermediate levels (fig. 19 A). Morphological changes were investigated by brightfield microscopy (fig. 19 A'). Cells differentiated for 7+3d in DM II displayed outgrowing neurites which led to the formation of a neuronal-like network after 7+14d of cultivation in DM II. At the later time point of differentiation, cell bodies clustered to aggregates (fig. 19 A').

Comparing **protocol IIa** with Ia only minor differences were detectable on transcript level (fig. 19 A and B). Expression levels of *P75NTR* did not decrease during differentiation as it was shown for protocol Ia. In addition, cells differentiated with protocol IIa showed increased mRNA expression values for *ASCL1* after 7+7d in DM III and for *GAP43* after 7+3d in DM III compared to the respective results generated with protocol Ia. For *TAC1*, a marginal expression was detectable in the 7+3d DM III specimen. Expression patterns of all other markers were comparable to the results described for protocol Ia (fig. 19 B). No obvious differences in the cell morphology at the analysed time points were observable if protocol IIa (fig. 19 B') instead of protocol Ia was used (fig. 19 A').

Transcriptional profiles of cells differentiated with **protocol IIIa** were highly comparable to the previously described ones for protocol Ia except for *TAC1* and *SYP* (fig. 19 A-C). Expression levels for the neuronal subtype specific marker *TAC1* increased gradually during *in vitro* differentiation although expression levels remained low. For the synaptic specific marker *SYP*, expression levels at later differentiation stages were slightly lower than in cells differentiated with protocol Ia or IIa (fig. 19 C). Based on brightfield microscopy analyses, this protocol was excluded from further investigations. Cells after 7+3d

in DM III of differentiation showed less and shorter neuronal-like outgrowths compared to differentiated cells of protocol Ia or IIa. This phenotype persisted as well at a later stage of neuronal maturation (7+14d in DM III) when a neuronal-like network was only marginally developed and floating cells indicative for apoptotic cells were detectable (fig. 19 C‘).

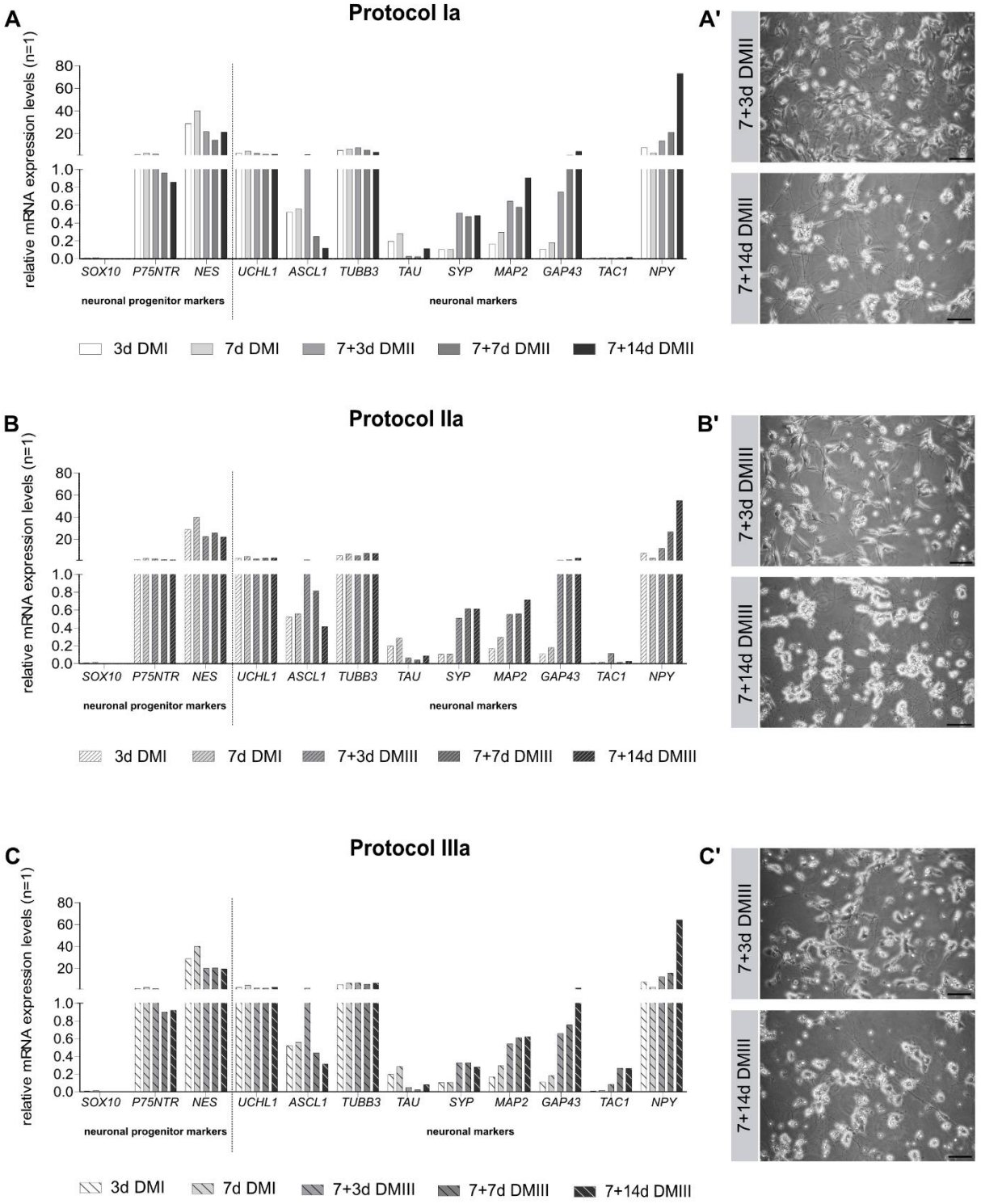


Figure 19: Results of pilot experiment A for the establishment of a neuronal differentiation protocol
A-C: Relative mRNA expression levels of differentiating SHSY5Y cells (protocol Ia (A), IIa (B), IIIa (C)) were assessed by qRT PCR using a set of different cell type markers. Column bar graphs show mean of two technical replicates (n=1). **A’-C’:** Cells were investigated for their morphology by brightfield microscopy after 7+3d or 7+14d under respective differentiating conditions. Scale bar: 75µm.

Pilot experiment B

In a second setup, protocol Ia and IIa were further modified. qRT PCR analyses were carried out for two biological replicates each and undifferentiated cells were investigated as well. Differentiation time was increased to up to 49 days in total.

Comparable to results of the first pilot experiment (A) (protocol Ia and IIa, fig. 19 A, B), no expression for the neuronal progenitor marker *SOX10* was detected using any of the experimental strategies (protocol Ib, IIb). Additionally, *TAC1* was as well only minimally expressed in respective differentiated cells (fig. 20 A, B).

SHSY5Y cells differentiated according to **protocol Ib** showed that relative mRNA expression levels for the neuronal progenitor marker *NES* slightly decreased during differentiation except for the latest investigated time point (7+42d in DM II). However, high standard deviations for this and other samples need to be taken into account. Contrastly to *NES,* the neuronal progenitor marker *P75NTR* was only marginally expressed in the cell population and decreased only during early phases of differentiation. After 7+14d DM II, *P75NTR* mRNA expression was upregulated again.
Comparable to results of the first pilot experiment, the pan-neuronal markers *UCHL1 and TUBB3* were almost stably expressed at intermediate levels throughout differentiation. The expression of *ASCL1* was strongly reduced comparing undifferentiated and 5d DM I differentiated samples. However, from 7+7d on, its expression was again strongly upregulated and then remained stable. In general, relative mRNA expression levels of *TAU, MAP2, SYP* and *GAP43* showed a gradual increase during differentiation although their overall expression levels were low to intermediate. *NPY* showed the highest mRNA expression levels among all neuronal markers. Neuronal cell fate induction caused first a decline in the transcript expression of this neuronal subtype specific gene, but further differentiation resulted in an expression upregulation (fig. 20 A).
Results of brightfield microscopy at early time points were highly comparable to the previously described findings of protocol Ia. Further differentiation led to the generation of a complex neuronal-like network after 7+28d DM II with cells displaying a high number of neuronal-like branches and synaptic-like contacts. The amount of dead cells floating in the cell culture dish was higher at later than at earlier stages of differentiation (fig. 20 A‘).

Results for relative mRNA expression analyses of cells differentiated with **protocol IIb** were highly comparable to results previously described for protocol Ib. Minor differences were detectable with regard to the mRNA expression levels of *P75NTR* as these were slightly higher compared to cells differentiated with protocol Ia. In contrast, the majority of the more advanced neuronal cell fate markers as *MAP2, SYP, GAP43* and *NPY* displayed slightly reduced expression levels especially at later differentiation time points compared to respective results of protocol Ib. Moreover, *ASCL1* expression was not stable in cells differentiated for longer periods but declined from 7+21d to 7+28d using protocol IIb (fig. 20 B).
In comparison to brightfield microscopy results of protocol IIa and protocol Ib, cell densities were considerably higher at all differentiation stages using this protocol. Additionally, the neuronal-like network was only marginally developed based on the presence of short neurite-like outgrowths after 7+28d of differentiation (fig. 20 B‘). Based on these observations, this protocol was excluded from further investigations and instead an optimized version of protocol Ib was selected for differentiation of all

genome edited and a *mock control* SHSY5Y clone. Specifically, the initial cell plating density and the cultivation time were modified (section 2.2.2).

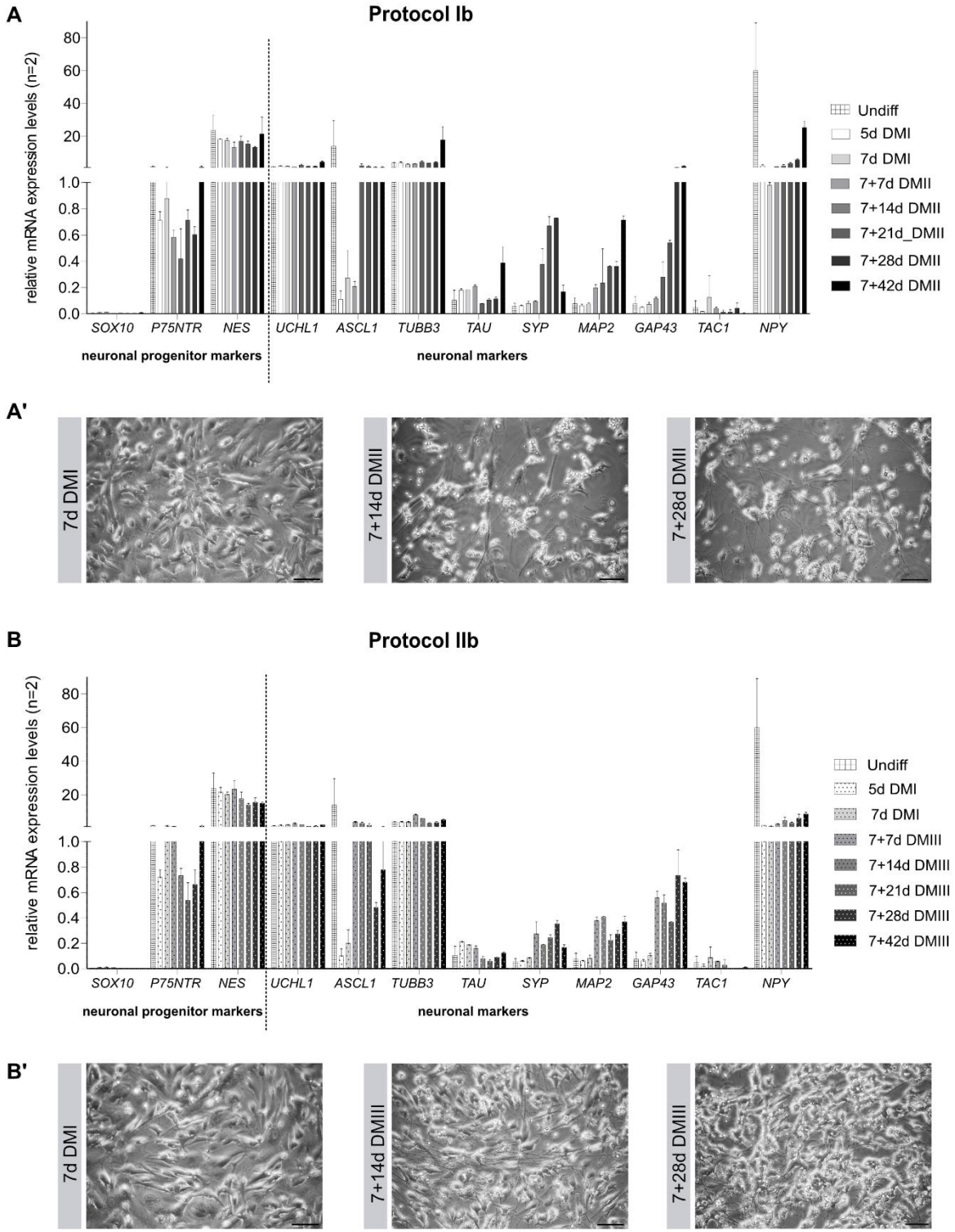


Figure 20: Results of pilot experiment B for the establishment of a neuronal differentiation protocol
A and B: Relative mRNA expression levels of undifferentiated and differentiating SHSY5Y (protocol Ib (A), IIb (B)) cells were assessed by qRT PCR using a set of different cell type markers. Bar plots show mean + standard deviation (STD) (n=2). **A' and B':** Cells were investigated for their morphology by brightfield microscopy after 7d, 7+14d and 7+28d of culture under respective conditions. Scale bar: 75µm.

3.3.3 *Neuronal differentiation of SHSY5Y clones*

Genome edited SHSY5Y clones were differentiated and assessed for neuronal maturation by brightfield microscopy. Morphological changes in the *mock control* clone during neuronal induction served as reference for observations.

Undifferentiated cells of the ***mock control*** clone showed a neuronal-like morphology with elongated cell bodies carrying several neurite-like outgrowths (fig. 21 A and A‘). Induction of neuronal maturation applying an optimized version of protocol Ib caused an accentuation of this phenotype after 7d of culture as neurite-like outgrowths further elongated (fig. 21 B and B‘). After 7+7d of differentiation, cell bodies started to accumulate in neuronal-like clusters (fig. 21 C and C‘). This phenotype became more pronounced after 7+14d of culture (fig. 21 D and D‘). From this stage on, neuronal-like clusters were highly organized and connected by semi adherent neurite-like connections. A slight increase in the clusters’ diameter was observable with proceeding differentiation time (fig. 21 D-F‘).

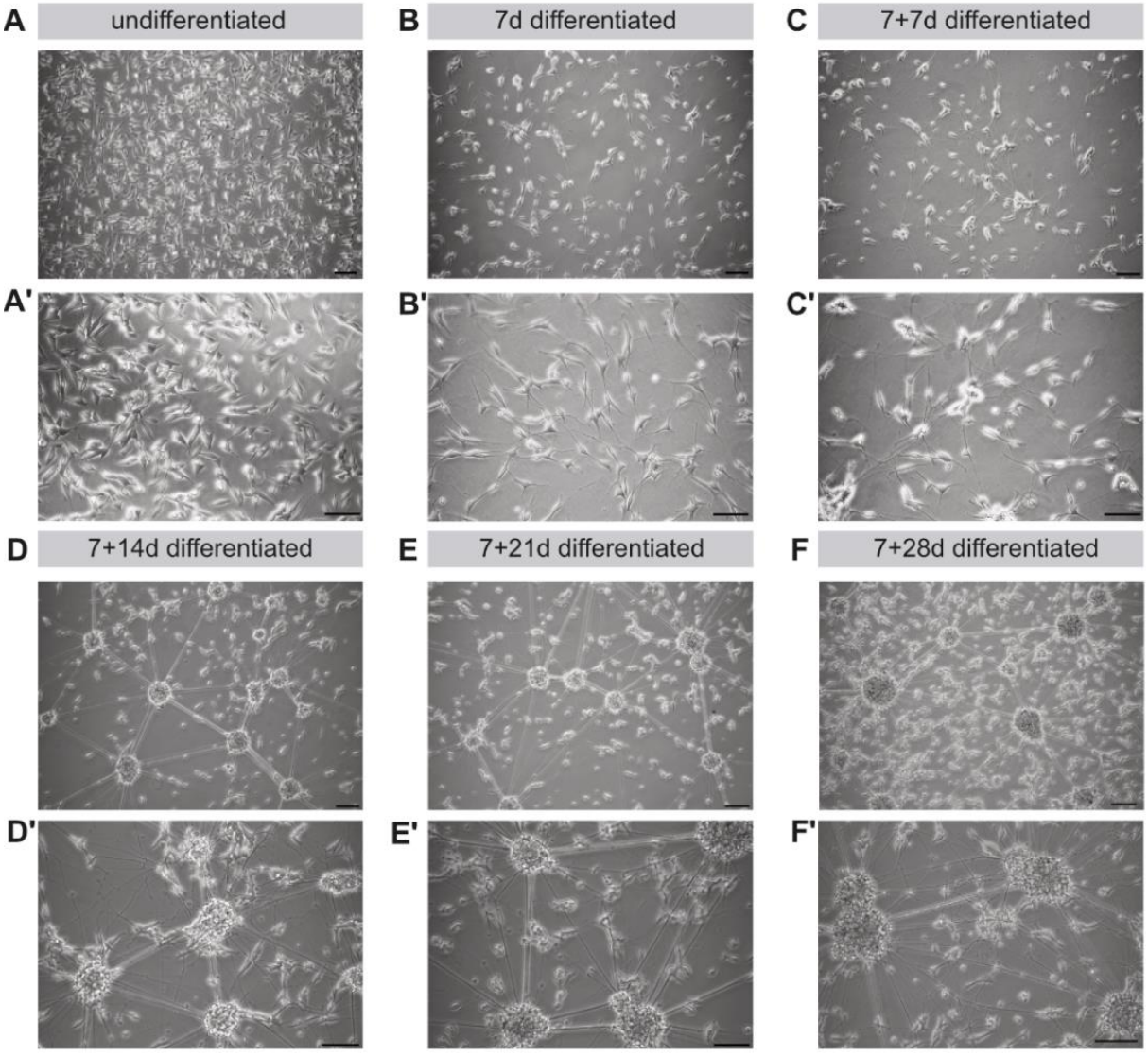


Figure 21: Neuronal-like maturation of the mock control clone
A-F’: Morphological changes in the cellular phenotype over the culturing period (undifferentiated (A, A’), 7d (B, B’), 7+7d (C, C’), 7+14d (D, D’), 7+21d (E, E’), 7+28d (F,F’) differentiated) were assessed by brightfield microscopy. Images are representatives of different experiments. Scale bar: A-F: 100µm, A’-F’: 75µm.

In the undifferentiated state, ***RET KO*** cells showed as well a neuronal-like morphology (fig. 22 A and A‘). Again, induction of neuronal maturation led to an elongation of cell bodies and neuronal-like outgrowths (fig. 22 B and B‘). A neuronal-like network started to form between 7d and 7+7d of differentiation when neuronal-like processes connected to each other (fig. 22 C and C‘). Neuronal-like clusters were first observable after 7+14d of differentiation but cellular organization was not as highly ordered as in the *mock control* clone (fig. 22 D and D‘). The number and diameter of clusters slightly increased with further cultivation time (fig. 22 E and E‘). After 7+28d, clusters were roundly shaped and cellular organization was comparable to observations for *mock control* cells (fig. 22 F and F‘).

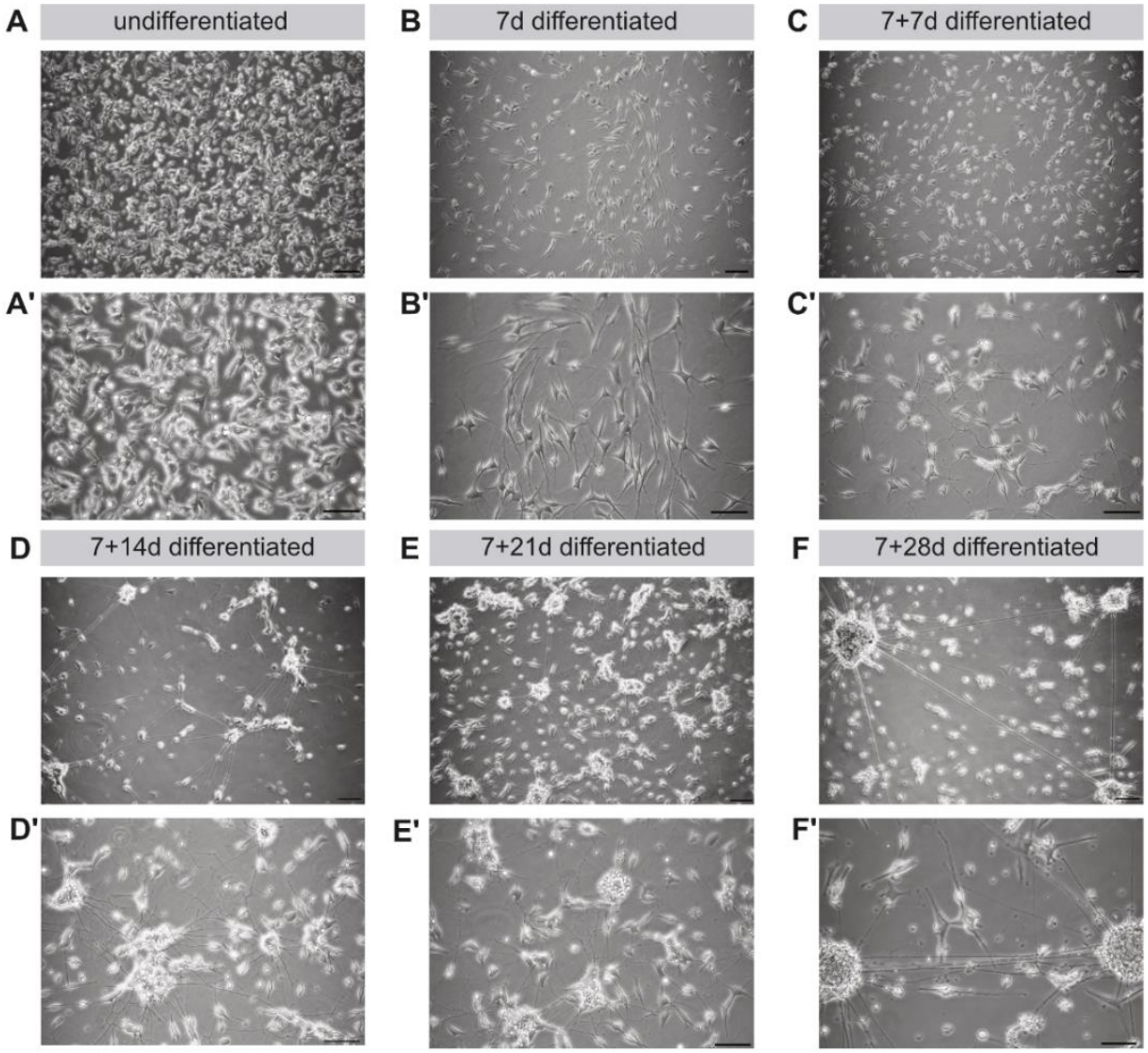


Figure 22: Differentiation of the RET KO clone into a neuronal-like phenotype
A-F': Brightfield microscopy images were taken at different time points of neuronal differentiation (undifferentiated (A, A'), 7d (B, B'), 7+7d (C, C'), 7+14d (D, D'), 7+21d (E, E'), 7+28d (F,F') differentiated). Images show representatives of different experiments. Scale bar: A-F: 100µm, A'-F': 75µm.

ATP7A KO cells displayed a neuronal-like morphology in their undifferentiated state. Besides an elongated cell body, individual cells presented already with neuronal-like outgrowths (fig. 23 A and A‘). In contrast to the *mock control* clone, *ATP7A* genome edited cells have formed a confluent monolayer after 7d of differentiation (fig. 23 B and B‘). An increase in the length of neuronal-like outgrowths was not observable before 7+7d of differentiation (fig. 23 C and C‘). The cell density was considerably higher at all time points of differentiation compared to the *mock control* clone (fig. 23 B-F‘). After 7+14d, a neuronal-like network was observable although cells tend to show less cell to cell connections than the *mock control* clone (fig. 23 D and D‘). Deviating to results obtained for the *mock control* clone, *ATP7A KO* cells did not show the formation of neuronal-like clusters at any stage of differentiation. In addition, the amount of presumably dead cells detaching from the culturing surface was higher - especially after 7+21d and 7+28d - compared to respective results of *mock control* clone (fig. 23 E-F‘).

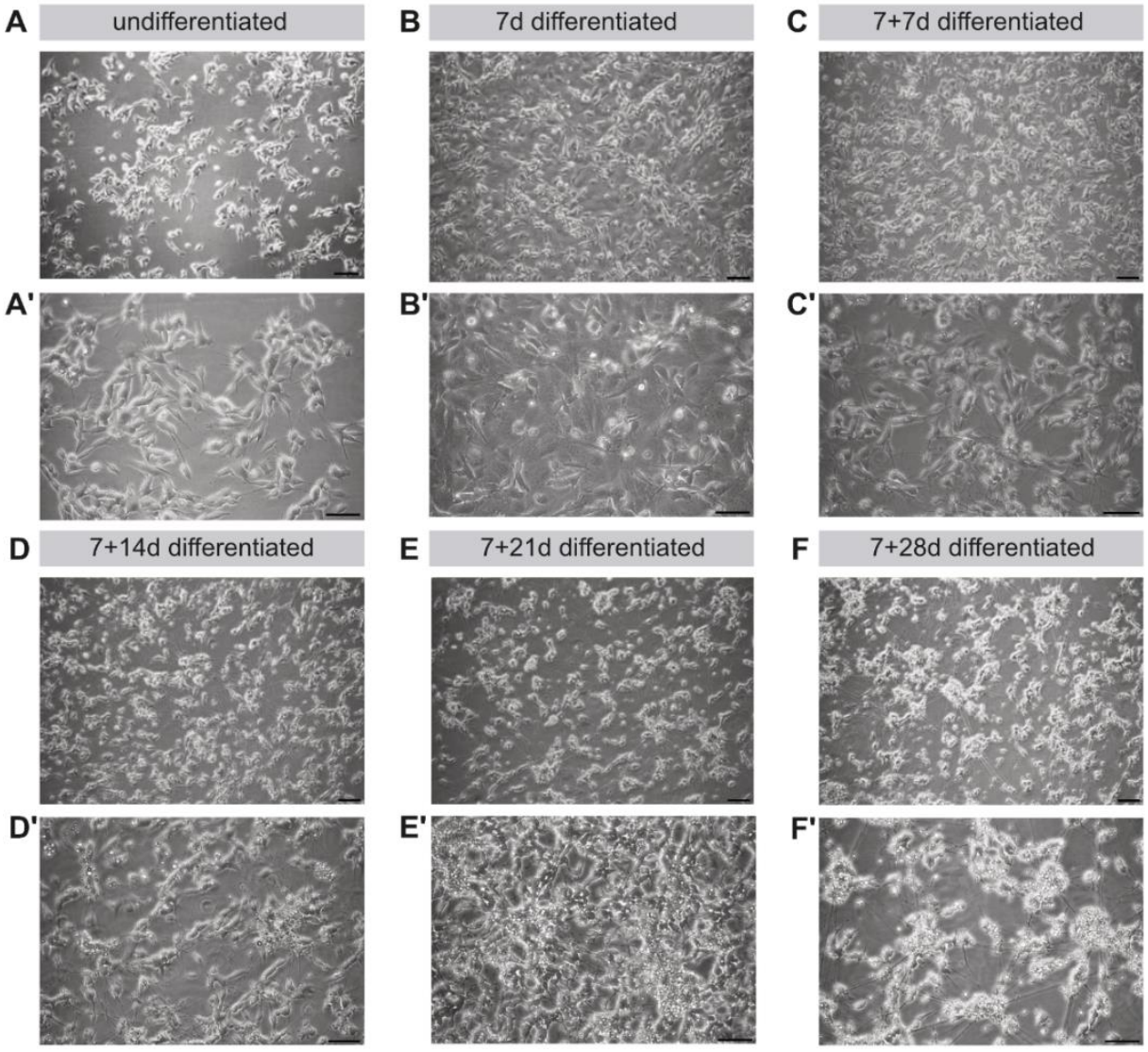


Figure 23: Neuronal-like cell fate maturation of the ATP7A KO clone
A-F': Cell morphologies of differentiating *ATP7A KO* cells were assessed by brightfield microscopy at various stages of neuronal differentiation (undifferentiated (A, A'), 7d (B, B'), 7+7d (C, C'), 7+14d (D, D'), 7+21d (E, E'), 7+28d (F,F') differentiated). Images display representatives of different experiments. Scale bar: A-F: 100µm, A'-F': 75µm.

Results of brightfield microscopy for the differentiating ***SREBF1 KO*** clone were highly comparable to observations made for the *RET KO* clone. Based on their cellular morphology, *SREBF1 KO* cells could be classified as neuronal-like in the undifferentiated state (fig. 24 A and A‘). The length of neurite-like outgrowths increased after 7d of differentiation (fig. 24 B and B‘). Further cultivation under differentiating conditions caused the formation of a neuronal-like network with increasing neuronal-tree like structures (fig. 24 C and C‘). Clustering of cell bodies started after 7+14d. A gain in the diameter and the organization of neuronal clusters was detectable at later stages of differentiation (fig. 24 D-F‘). However, the arrangement of neuronal-like clusters was not as highly ordered as in the *mock control* clone at any of the investigated time points.

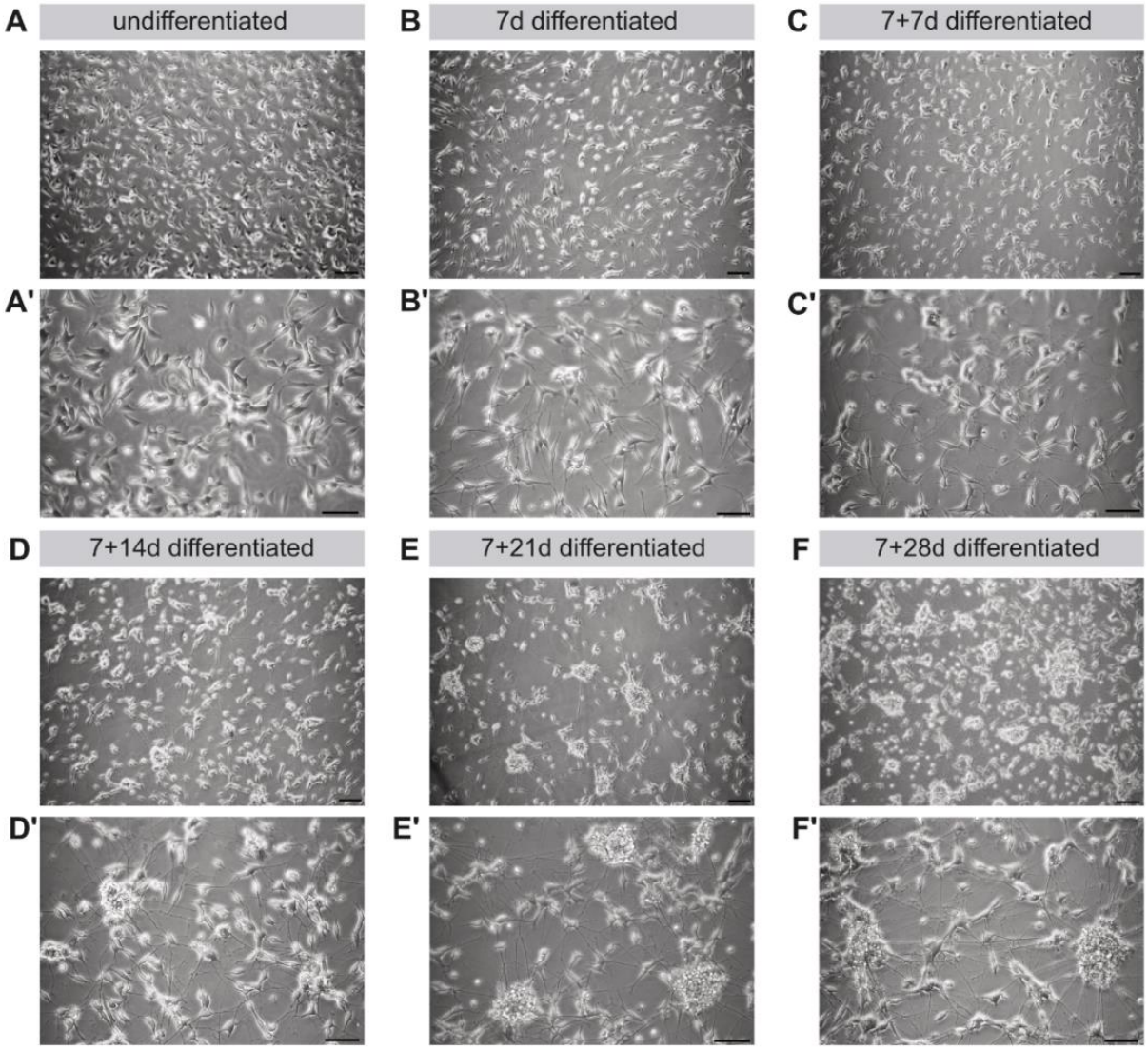


Figure 24: Neuronal-like cell fate induction in the SREBF1 KO clone
A-F': Brightfield microscopic analyses were performed to monitor the cellular morphology during neuronal differentiation (undifferentiated (A, A'), 7d (B, B'), 7+7d (C, C'), 7+14d (D, D'), 7+21d (E, E'), 7+28d (F,F') differentiated). Images are representatives of different experiments. Scale bar: A-F: 100µm, A'-F': 75µm.

Neuronal-like cells were as well observable for the undifferentiated ***ABCD1 KO*** clone (fig. 25 A and A‘). An interconnection between individual cells via neuronal-like outgrowths started to form after 7d of differentiation (fig. 25 B and B‘). Neurite-like structures extended in length and cell bodies aggregated in neuronal-like clusters with further cultivation time (fig. 25 C-D‘). After 7+21d of differentiation, the cellular pattern of *ABCD1 KO* cells was highly comparable to the organization of *mock control* cells at the respective differentiation stage. Neuronal clusters were highly structured, roundly shaped and connected by semi-adherent neurite-like connections (fig. 25 E-F‘).

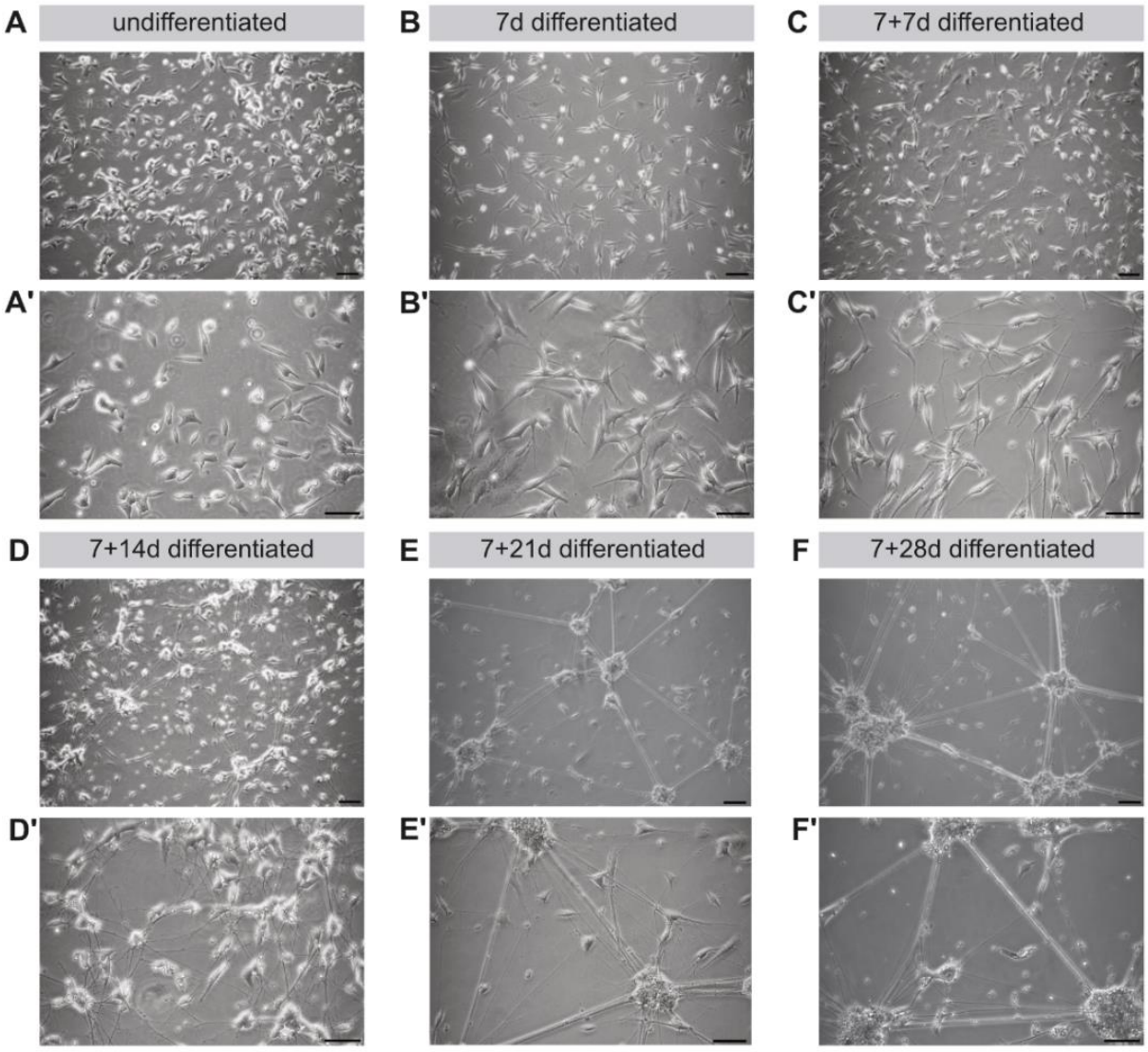


Figure 25: Differentiation of the ABCD1 KO clone into neuronal-like cells
A-F': Differentiating cells were investigated by brightfield microscopy at different time points of cultivation (undifferentiated (A, A'), 7d (B, B'), 7+7d (C, C'), 7+14d (D, D'), 7+21d (E, E'), 7+28d (F,F') differentiated). Images reflect representatives of different experiments. Scale bar: A-F: 100µm, A'-F': 75µm.

3.3.4 *Expression analyses of SHSY5Y clones*

Besides the morphological assessment, all genome edited clones as well as the *mock control* clone were further investigated on transcript and protein level (fig. 26-30).
In general, the progenitor marker *SOX10,* and the glial marker *(GLIAL FIBRILLARY ACIDIC PROTEIN) GFAP* were not expressed on mRNA level in any of the clones. An exploratory data analysis was carried out for qRT PCR expression data of all *KO* clones in order to to obtain first indications about putatively deregulated markers. To this end, the mean relative mRNA expression level of a marker at a specific differentiation stage was compared between the *mock control* and a specific *KO* clone using a two-sided unpaired T-Test (with/without Welch's correction). Table 8 provides an overview about all differentially expressed genes in the individual gene-specific *KO* clones at specific time points of differentiation. However, respective data can therefore just be interpreted on descriptive level. These significant, quantitative discrepancies between the individual *KO* clones in comparison to the *mock control* clone will not be addressed in the description of the following figures (fig. 26-30 A). In contrast, only general differences in the overall marker expression profiles comparing the *KO* clones and the *mock control* clone will be described.
A subset of neuronal markers was selected for expression validation on protein level by IF analyses. Specifically, to assess proceeding neuritogenesis and neuronal maturation at 7+14d and 7+28d of differentiation, the cytoskeleton proteins TUBB3, MAP2, TAU, the growth cone specific protein GAP43 and the synaptic marker SYP were chosen. To study the proliferative behaviour of the differentiating clones, the proliferation marker ANTIGEN Ki-67 (Ki67) was included into the analyses (fig. 26-30 B and C).

In the ***mock control*** clone, expression levels of the progenitor marker *P75NTR* increased until 7+7d, whereas with progressing differentiation, a gradual decrease was observed. The progenitor marker *NES* showed higher expression levels than *P75NTR* but a similar expression pattern. The majority of neuronal cell fate markers, such as *UCHL1, TUBB3, MAP2, SYP, GAP43* and *NPY* showed a clear upregulation of expression after induction of neuronal differentiation at 7d in DM I. Compared to expression levels of *UCHL1*, *TUBB3, GAP43* and *NPY*, the markers *MAP2* and *SYP* displayed considerably lower expression in differentiated cells. *TAU* expression was generally low but gradually increased during differentiation. In contrast, *ASCL1* levels first strongly declined from undifferentiated to 7d differentiated cells and then remained almost stable during further cultivation. The neuronal subtype specific marker *TAC1* was expressed at low levels in undifferentiated and 7d differentiated cells, while samples at later differentiation stage showed higher but highly variable expression levels based on respective standard errors. The apoptotic marker *CASPASE 3 (CASP3)* displayed increased but stable expression levels in differentiated samples compared to the undifferentiated specimens (fig. 26 A).
Expression of selected neuronal cell fate markers could be confirmed by IF analyses. Corresponding to their respective protein functions, TUBB3, MAP2, GAP43 and TAU localised to neuronal-like outgrowths and showed a partial expression overlap in corresponding IF double staining. The synaptic specific marker SYP generated a dotted-like staining pattern along TAU^+ neurite-like structures. IF staining of Ki67 demonstrated that only a subfraction of $TUBB3^+$ cells remained mitotically active after 7+14d and 7+28d of differentiation, respectively (fig. 26 B and C). Corresponding to brightfield microscopy results,

IF analyses of cells after 7+14d and 7+28d displayed proceeding neuronal-like maturation (fig. 21, 26 B and C).

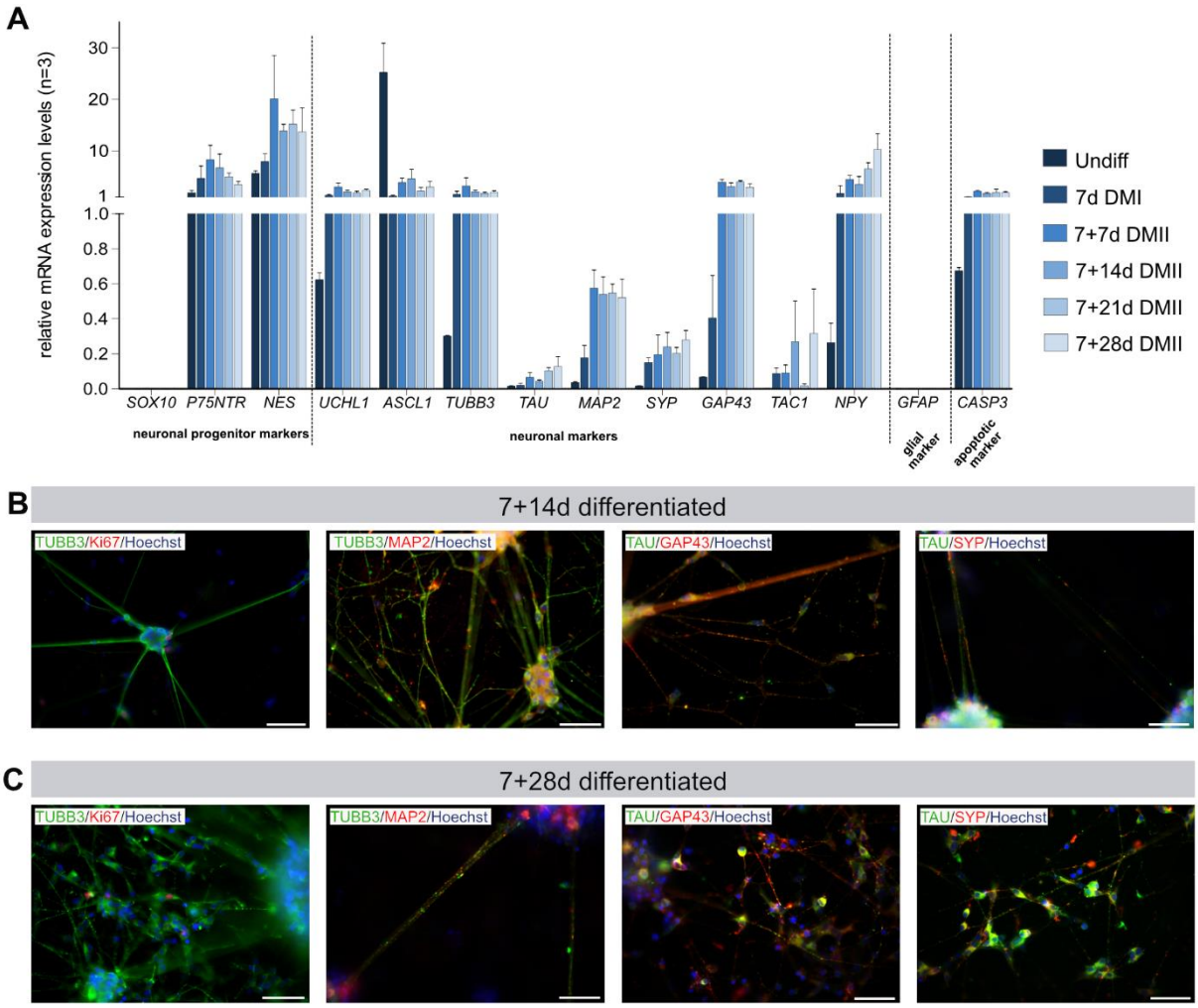


Figure 26: Expression analyses of the differentiating mock control clone
A: mRNA expression profiles were assessed at various time points of neuronal differentiation by qRT PCR (n=3, mean + SEM (standard error of mean). **B and C:** Differentiated cells after 7+14d (B) and 7+28d (C) of cultivation were investigated by IF analyses for selected marker expression. Images are representatives of different experiments. Cell type specific markers are indicated in red (neuronal: MAP2, GAP43, SYP; proliferation: Ki67) and green (TUBB3, TAU). Nuclei were counterstained with Hoechst (blue). Scale bar: 50µm.

Overall, results of mRNA expression profiling were highly comparable between the *mock control* clone and the ***RET KO*** clone (fig. 26 A and 27 A). Significant differences in the relative expression levels of all markers comparing both clones on descriptive level are summarized in table 8. Major differences in the general expression profile of individual markers are described in the paragraph below.

The relative mRNA expression profile of the progenitor marker *P75NTR* varied, comparing both clones. Specifically, a strong upregulation in the relative *P75NTR* mRNA expression was observable comparing undifferentiated and 7d differentiated *RET KO* cells. During further differentiation, transcript levels of this marker slightly increased. Minor differences were as well detectable for *NES*. Comparable to the *mock control* clone, expression of this progenitor marker increased during early stages of differentiation but then decreased after 7+21d in the *RET KO* clone. Comparable observations were made for the neuronal markers *SYP* and *GAP43* as marker expression levels decreased as well in the latest differentiation

specimen in the *RET KO* clone but not in the *mock control* clone. The pan-neuronal marker *UCHL1* displayed a strong decline in the relative mRNA expression comparing undifferentiated and 7d differentiated cells while further cultivation caused again an increase. For *TAC1* only the 7d DMI sample displayed a minimal mRNA expression. The relative mRNA expression of the apoptotic marker *CASP3* increased just after 7d of differentiation and then remained stable. Expression patterns for all other markers were comparable as described for the *mock control* clone (fig. 27 A).

Results of IF analyses were as well highly comparable between the *mock control* clone and the *RET KO* clone. Marker staining validated an expression of the neuronal cytoskeleton proteins TUBB3, MAP2 and TAU as well as of the growth cone specific protein GAP43 in differentiating *RET KO* cells. Staining patterns indicated neuronal-like morphologies after 7+14d and 7+28d of differentiation. SYP signals displayed the presence of synaptic-like contacts scattered along the neurite-like structures immunoreactive for TAU. Comparable to the *mock control* clone, only a minority of differentiating cells were positive for the proliferation marker Ki67 (fig. 27 B and C).

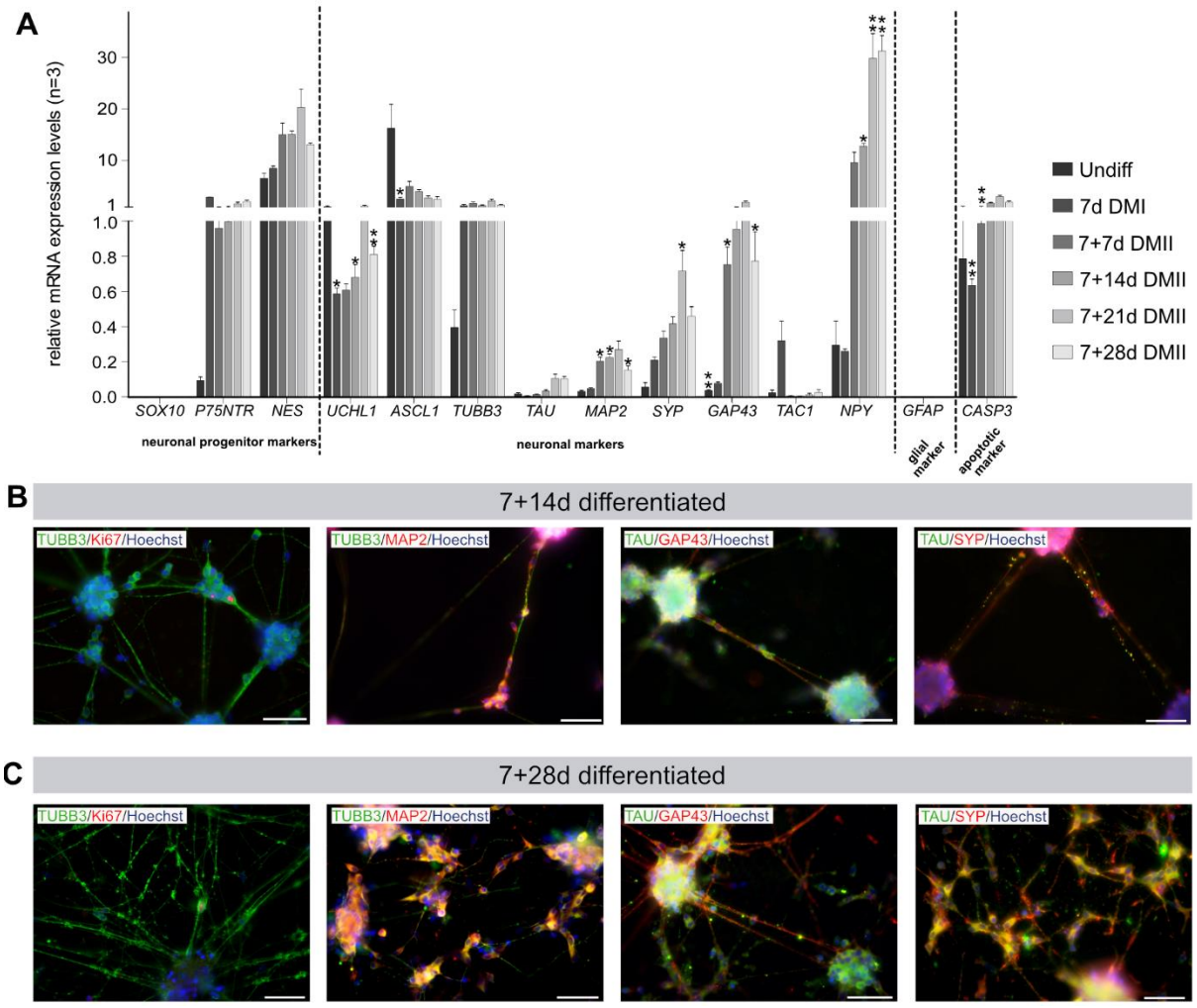


Figure 27: Marker gene expression in the differentiating RET KO clone

A: qRT PCR analyses using *RET KO* cells of various maturation states (n=3, mean + standard error of mean (SEM), two-sided unpaired T-Test with/without Welch's correction *p<0.05; ** p<0.01; *** p< 0.001). **B and C:** Cells after 7+14d (B) and 7+28d (C) of differentiation were investigated by IF analyses for marker expression patterns. Images are representatives of different experiments. Cell type specific markers are indicated in red (neuronal: MAP2, GAP43, SYP; proliferation: Ki67) and green (TUBB3, TAU). Nuclei were counterstained with Hoechst (blue). Scale bar: 50µm.

Exploratory data analyses of mRNA expression profiles between the *mock control* clone and the ***ATP7A KO*** clone revealed some differences (fig. 26 A and 28 A). Again, all differentially expressed genes and the respective time points are summarized in table 8.

In contrast to the *mock control* clone, mRNA expression levels for both neuronal progenitor markers, *P75NTR* and *NES*, remained almost stable throughout differentiation, except for stage 7+21d that showed upregulated transcript levels. Expression of the pan-neuronal marker *UCHL1* increased only during later stages of *in vitro* differentiation while expression levels of *TUBB3* were stable comparing undifferentiated and all differentiated specimens. In contrast to the *mock control* clone, mRNA expression levels of *TAU* were higher in the undifferentiated and 7d differentiated samples than in cells differentiated for longer periods. Moreover, *TAC1* expression was not detectable in any specimen. mRNA levels of the neuronal subtype specific marker *NPY* first declined from the undifferentiated to the 7+7d differentiated state before increasing at later stages of differentiation. The apoptotic marker *CASP3* was almost stably expressed at low levels in all samples (fig. 28 A).

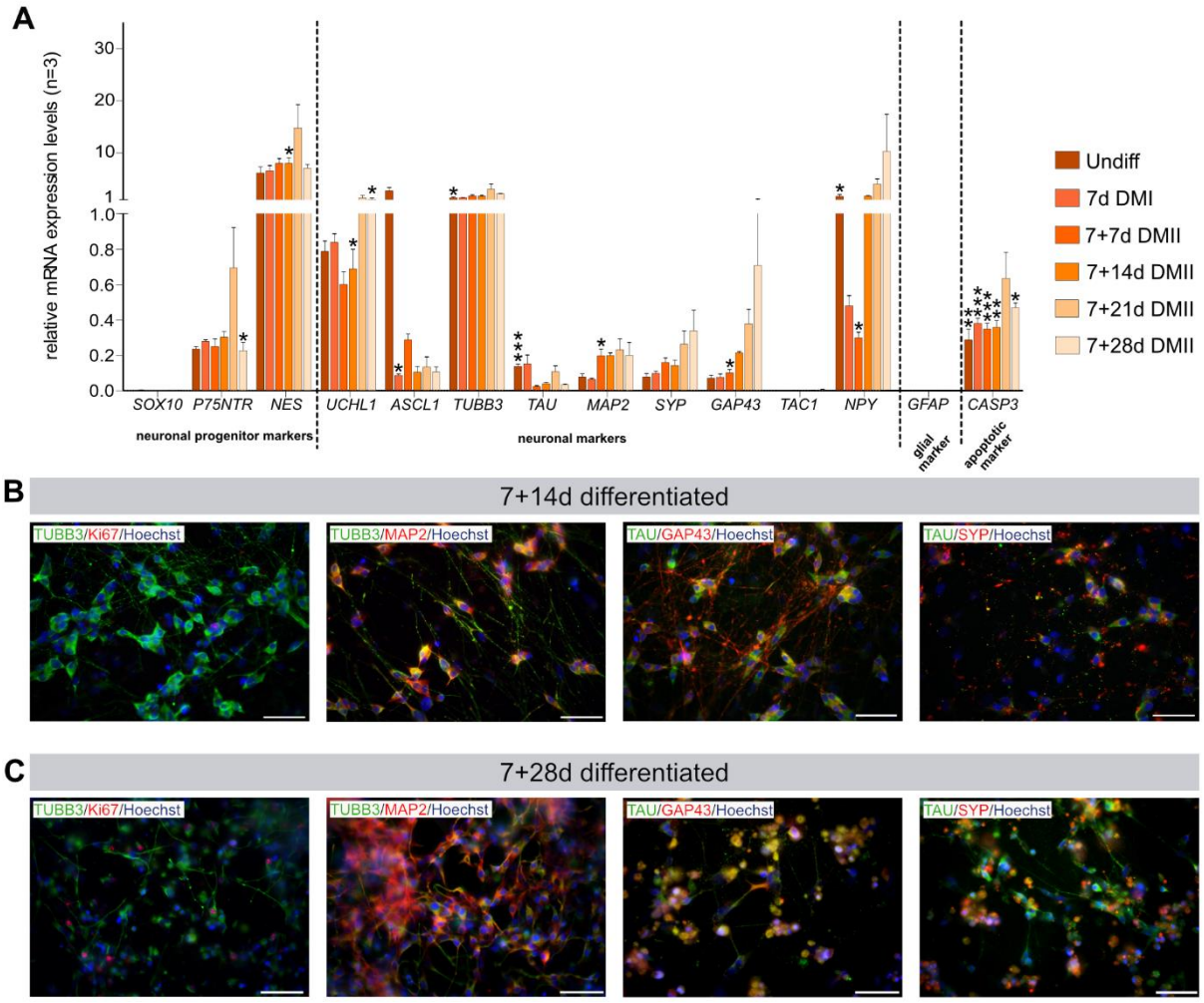


Figure 28: Marker expression profiling in the differentiating ATP7A KO clone

A: mRNA expression profiling by qRT PCR using undifferentiated and differentiated *ATP7A KO* cells of various differentiation states (n=3, mean + standard error of mean (SEM), two-sided unpaired T-Test with/without Welch's correction *p<0.05; ** p<0.01; *** p< 0.001). **B and C:** 7+14d (B) and 7+28d (C) differentiated cells were investigated for marker expression by IF analyses. Images are representatives of different experiments. Cell type specific markers are indicated in red (neuronal: MAP2, GAP43, SYP; proliferation: Ki67) and green (TUBB3, TAU). Nuclei were counterstained with Hoechst (blue). Scale bar: 50µm.

IF analyses did validate respective neuronal marker expression for TUBB3, MAP2, GAP43 and SYP after 7+14d of differentiation. In contrast to the *mock control* clone, no neuronal-like clusters were observable, and the axonal marker TAU did not generate a distinct staining pattern at this stage (fig. 28 B). While after 7+14d only a minority of cells was immunoreactive for the proliferation marker Ki67, after 7+28d the number of this subpopulation was elevated. Generally, at this time point of differentiation, the overall cell density was high and a neuronal-like network was only poorly developed. This caused problems during cell fixation and IF analyses as cells easily detached from the surface, which was observable in IF analyses of TAU/GAP43 and TAU/SYP. Consequently, TAU, GAP43 and SYP specific signals were hardly observable in these samples (fig. 28 C).

Overall, relative mRNA expression profiles of the ***SREBF1 KO*** were highly comparable to results obtained for the *mock control* clone although a number of markers was differentially expressed at specific time points of neuronal differentiation (tab. 8, fig. 26 A and 29 A). This was revealed by an exploratory data analysis strategy.

Expression patterns of *NES, TAU, TUBB3, SYP, GAP43* and *NPY* were highly comparable to results described for the *mock control* clone as they showed generally upregulated expression levels comparing undifferentiated and differentiated specimens. In comparison to the *mock control* clone, transcript levels of the neuronal progenitor marker *P75NTR* increased drastically from undifferentiated to 7d differentiated *SREBF1 KO* cells. Additionally, the pan-neuronal marker *UCHL1* and the apoptotic marker *CASP3* showed higher expression levels in differentiated than in undifferentiated samples while stable relative transcript levels for both genes were reached only after 7+7d in this *KO* clone. The downregulation in the relative expression for *ASCL1* comparing undifferentiated and differentiated *SREBF1 KO* cells was not as pronounced as in the *mock control* clone. Moreover, although *MAP2* increased during early phases of differentiation, this upregulation was not maintained as well at later stages. For *TAC1,* all samples displayed only a minimal expression without any regulation (fig. 29 A).

On protein level, results of IF analyses confirmed respective neuronal marker expression in the differentiated *SREBF1 KO* after 7+14d and 7+28d of cultivation, respectively. In general, proceeding neuronal maturation was observable as e.g. TAU specific signals were more prominent after 7+28d than after 7+14d. Staining patterns e.g. in neurite-like structures were in line with respective protein functions as in the case of the cytoskeletal protein TUBB3, the dendrite marker MAP2 and the axonal marker TAU. Again, comparable to the *mock control* clone only a minority of cells remained proliferatively active as it was observable based on Ki67 specific nuclear signals (fig. 29 B and C).

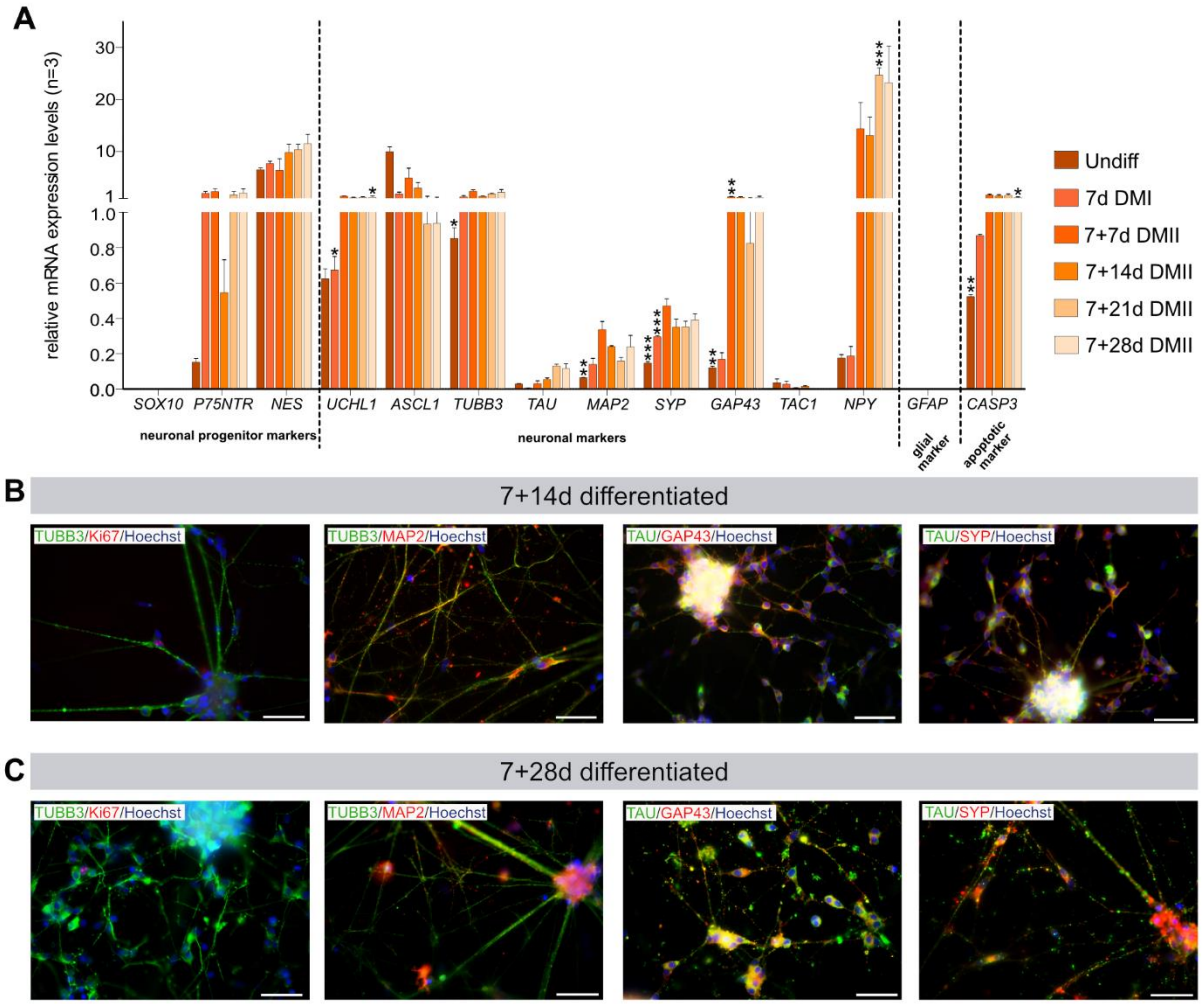


Figure 29: Expression profiling of the SREBF1 KO during neuronal in vitro differentiation
A: Expression analyses of undifferentiated and differentiated *SREBF1 KO* cells assessed by qRT PCR. Several neuronal differentiation states were analyzed (n=3, mean + standard error of mean (SEM), two-sided unpaired T-Test with/without Welch's correction *p<0.05; ** p<0.01; *** p< 0.001). **B and C:** Differentiated cells (7+14d (B) and 7+28d (C)) were fluorescently labelled with a subset of selected markers. Images are representatives of different experiments. Cell type specific markers are indicated in red (neuronal: MAP2, GAP43, SYP; proliferation: Ki67) and green (TUBB3, TAU). Nuclei were counterstained with Hoechst (blue). Scale bar: 50 µm.

Expression profiling results of the ***ABCD1 KO*** clone displayed only minor differences in comparison to the *mock control* clone (fig. 26 A and 30 A). In table 8, all differentially expressed genes and the corresponding time points of differentiation, which were identified by exploratory data analysis, are summarized.

Differences between both clones were observable for the two investigated neuronal progenitor markers *P75NTR* and *NES*, as expression levels of *P75NTR* slightly increased during differentiation, while relative transcript levels of *NES* decreased in the context of differentiation. Moreover, in contrast to the *mock control* clone, *UCHL1* did not display any upregulation during *in vitro* differentiation and *TUBB3* transcript levels only increased at later stages of differentiation. Comparable to the *mock control* clone, the expression of most of the advanced neuronal cell fate markers as *MAP2, SYP* and *GAP43* gradually increased in the context of differentiation but - deviating to the *mock* specific results - reached stable and most prominent expression levels already at 7+14d to 7+21d. Another striking difference comparing

both clones was detectable for *NPY*, as this neuronal subtype specific marker first decreased directly after neuronal cell fate induction before increasing with progressive differentiation time. Expression patterns for *ASCL1* and *TAU* in differentiated *ABCD1 KO* samples were comparable to the results described for the *mock control* clone. Stable but low or minimal expression levels were detectable for the apoptotic marker *CASP3* and the neuronal subtype specific marker *TAC1,* respectively (fig. 30 A). Neuronal marker expression on protein level was validated for TUBB3, MAP2, TAU, SYP and GAP43 and respective staining patterns confirmed expected subcellular protein localisation. Proceeding neuronal-like maturation was observable based on larger neuronal-like clusters and more complex neuronal-like networks comparing cells after 7+14d and 7+28d of differentiation. Similar to the *mock control* clone, only a few TUBB3$^+$ cells displayed immunoreactivity for Ki67 (fig. 30 B and C).

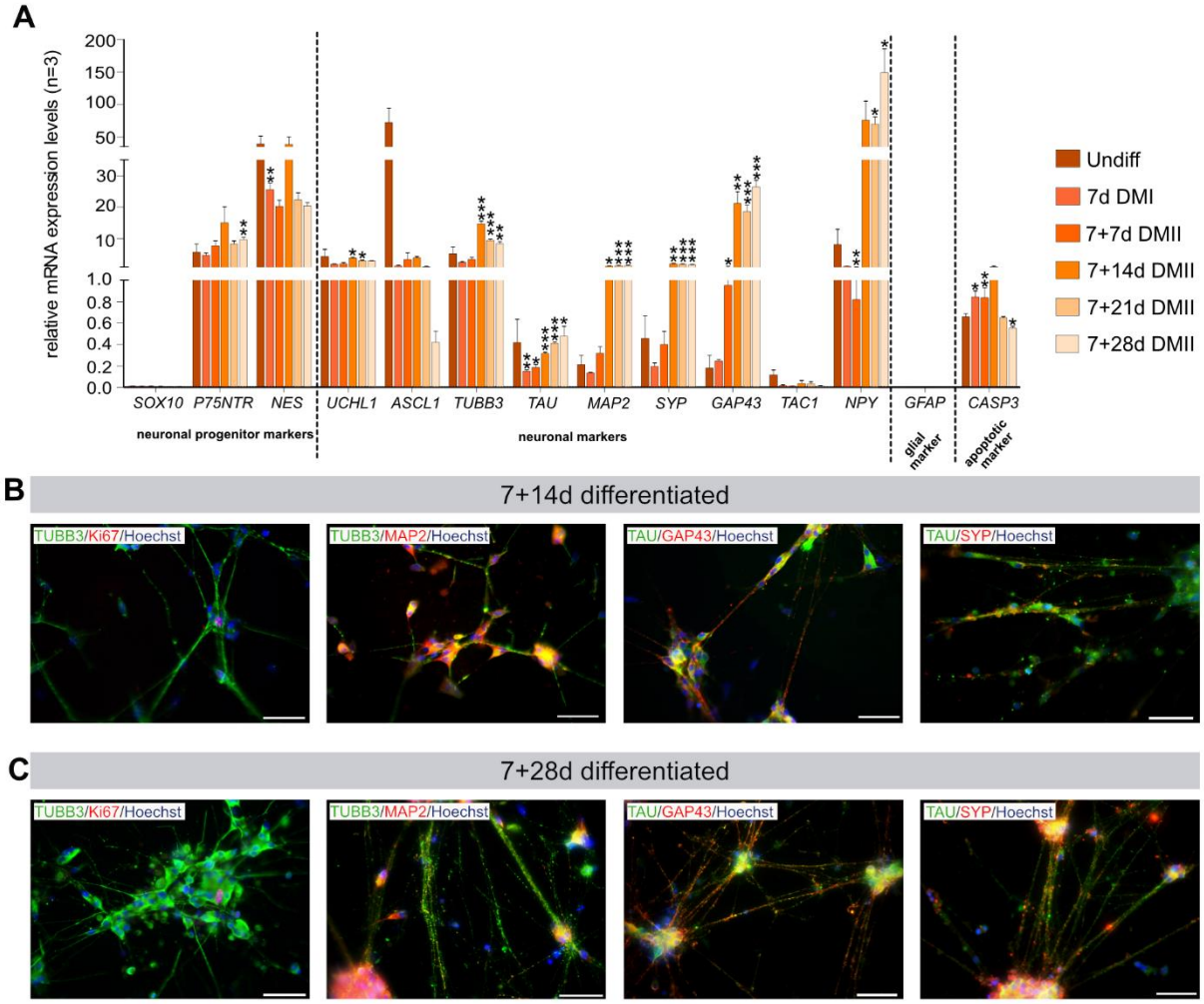


Figure 30: Expression profiling of the differentiating ABCD1 KO

A: qRT PCR investigations of undifferentiated and differentiated *ABCD1 KO* cells. Various neuronal maturation states were analysed (n=3, mean + standard error of mean (SEM), two-sided unpaired T-Test with/without Welch's correction *p<0.05; ** p<0.01; *** p< 0.001). **B and C:** 7+14d (B) and 7+28d (C) differentiated cells were analysed by IF staining for the expression of selected markers. Images are representatives of different experiments. Cell type specific markers are indicated in red (neuronal: MAP2, GAP43, SYP; proliferation: Ki67) and green (TUBB3, TAU). Nuclei were counterstained with Hoechst (blue). Scale bar: 50µm.

Besides the presented cell type specific markers, mRNA expression profiles of the selected HSCR candidate genes (*ATP7A, SREBF1* and *ABCD1*) and the major HSCR risk locus *RET* were assessed as it was aimed to determine if candidate genes might act up- or downstream of the RET signalling pathway. Therefore, only the *mock control* clone and the *RET KO* clone were analysed for the expression of all previously listed genes, while all candidate specific clones were only investigated for *RET* and the edited loci.

In the *mock control* clone, transcript levels of all investigated genes strongly increased after 7d of neuronal induction and then decreased gradually with further differentiation. In general, *ABCD1* relative expression levels were lowest (fig. 31 A).

Candidate gene-specific results were comparable between the *mock control* clone and the *RET KO* clone. Contrastly to these similarities, *RET* expression levels increased at stage 7+21d in the *RET KO* clone which was not observable in the *mock control* clone (fig. 31 B).

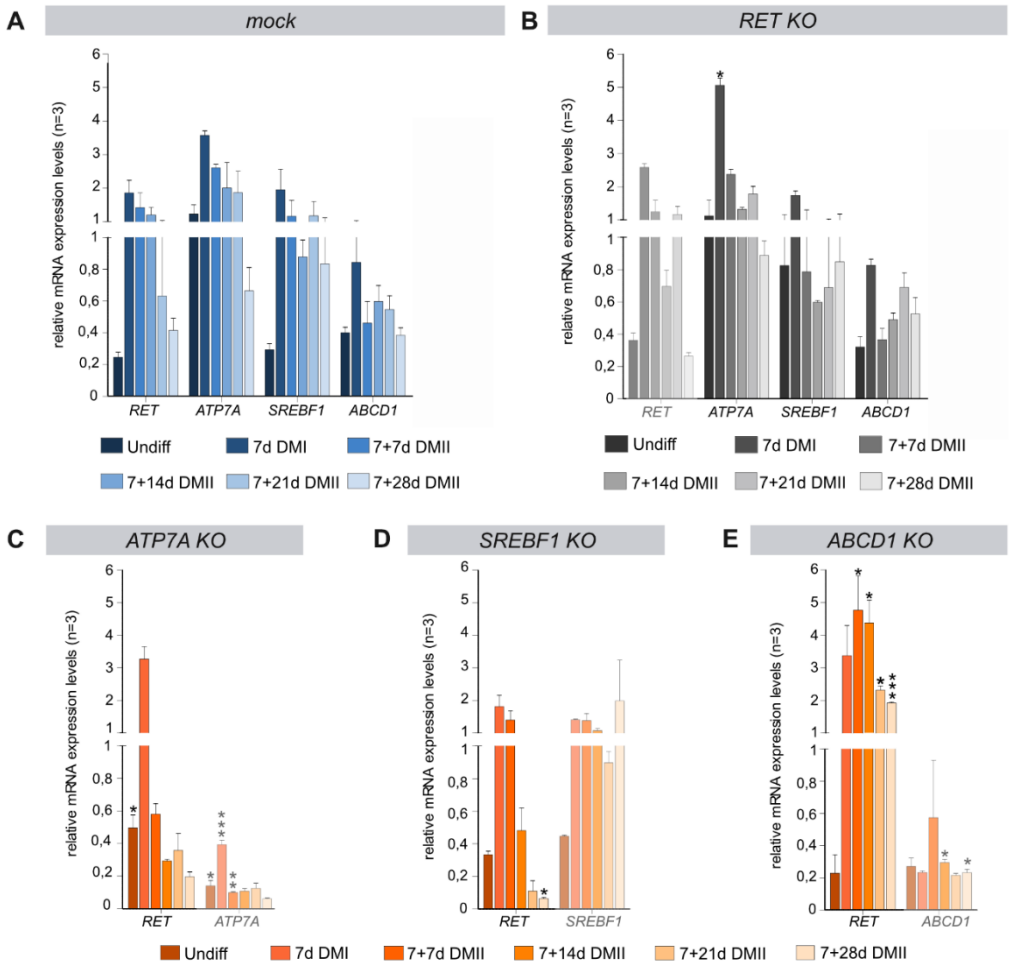


Figure 31: mRNA expression profiling of HSCR candidate genes and RET in individual clones

A and B: qRT PCR investigations of the *mock control* (A) and the *RET KO* (B) clones for *RET* and the HSCR candidate genes *ATP7A*, *SREBF1* and *ABCD1*. **C-E:** Transcriptional profiling of *RET* and the respective genome edited loci in *ATP7A KO* (C), *SREBF1 KO* (D) and *ABCD1 KO* (E) clones. Undifferentiated and differentiated cells of various neuronal differentiation stages were analysed by qRT PCR, respectively (n=3, mean + standard error of mean (SEM), two-sided unpaired T-Test with/without Welch's correction. *p<0.05; ** p<0.01; *** p< 0.001). Results for the genome edited loci of the corresponding *KO* clone are faded.

Although a strong upregulation in the expression upon neuronal induction of *RET* was observable in all three *KO* clones, while expression patterns at later differentiation stages varied between the clones and the *mock control* clone. Specifically, in the *ATP7A KO* clone, respective gene expression decreased again after the initial upregulation and then remained almost stable in all differentiated specimens (fig. 31 C). For the *SREBF1 KO* clone, *RET* expression was almost stable throughout differentiation while for the *ABCD1 KO* clone the respective expression gradually diminished after 7+7d (fig. 31 D, E).
Differences in the transcript expression profiles were as well detectable for the genome edited loci in the respective candidate specific *KO* clone in comparison to the *mock control* clone. Specifically, *ATP7A* relative expression levels in the *ATP7A KO* clone were generally low and stable - except for the 7d sample. Relative transcript levels of *SREBF1* in the *SREBF1 KO* clone increased after 7d and then remained constant; respective levels of *ABCD1* in the corresponding KO clone were as well low but almost stable in all samples (fig. 31 C-E).

Results of the exploratory mRNA expression data analyses are summarized for all cell type specific markers and *RET* in table 8. Specifically, changes in expression profiles (up- or downregulation) which were revealed by the descriptive statistical data evaluation comparing respective *KO* clones to the *mock control* clone, are indicated. As it can be seen, in all gene-specific *KO* clones a variety of cell type markers was detected to be differentially expressed at individual stages of differentiation. The type of deregulation (up-/downregulation) for a specific gene was mostly stable throughout differentiation in the respective clone (tab. 8).

Table 8: Overview on the comparative expression analysis of individual KO clones at different time points of differentiation

Only significant changes (exploratory data analysis using a two-sided unpaired T-Test) in expression levels of gene-specific *KO* clones in comparison to the *mock control* clone are displayed. Up- or downregulation at the investigated timepoint is illustrated schematically (↓downregulation/↑upregulation). Relative mRNA expression levels of *SOX10* and *GFAP* were not expressed and therefore not taken into account. Undiff (Undifferentiated)

Relevance	Marker	*RET KO*	*ATP7A KO*	*SREBF1 KO*	*ABCD1 KO*
Neuronal progenitors	*SOX10*	not expressed			
	P75NTR				
	NES				
Immature/ Mature neurons	*UCHL1*	7d ↓, 7+14d ↓, 7+28d ↓	7d ↓, 7+14d ↓, 7+28d ↓	7d ↓, 7+14d ↓, 7+28d ↓	7d ↓, 7+14d ↓, 7+28d ↓
	ASCL1	7d ↑	7d ↑	7d ↑	7d ↑
	TUBB3				
	TAU				
	MAP2	7+7d ↓, 7+14d ↓, 7+28d ↓	7+7d ↓, 7+14d ↓, 7+28d ↓	7+7d ↓, 7+14d ↓, 7+28d ↓	7+7d ↓, 7+14d ↓, 7+28d ↓
	SYP	7+21d ↑	7+21d ↑	7+21d ↑	7+21d ↑
	GAP43	Undiff ↓, 7+7d ↓, 7+28d ↓	Undiff ↓, 7+7d ↓, 7+28d ↓	Undiff ↓, 7+7d ↓, 7+28d ↓	Undiff ↓, 7+7d ↓, 7+28d ↓
	TAC1				
	NPY	7+14d ↑, 7+21d ↑, 7+28d ↑	7+14d ↑, 7+21d ↑, 7+28d ↑	7+14d ↑, 7+21d ↑, 7+28d ↑	7+14d ↑, 7+21d ↑, 7+28d ↑
Glia	*GFAP*	not expressed			
Apoptosis	*CASP3*	7d ↓, 7+7d ↓	7d ↓, 7+7d ↓	7d ↓, 7+7d ↓	7d ↓, 7+7d ↓
Major HSCR risk gene	*RET*				

3.3.5 *Functional analyses of SHSY5Y clones*

Genome edited clones were characterized regarding functional impairments in comparative analyses with the *mock control* clone. *In vitro* assays targeting cell proliferation, cell survival and cell migration capacity were performed. Besides the undifferentiated clones, differentiated clones of stage 7+1d or 7+7d were investigated. These stages were chosen due to specific experimental based reasons as outlined below or as they were regarded as most informative. The fact that cell clones tend to form extensive cell clusters at later differentiation time points made it impossible to reliably determine proliferative or apoptotic percentages by microscopic image analyses.

Cell proliferation

Cell proliferation rates of undifferentiated and 7+1d differentiated cells were determined by BrdU assays (fig. 32 A). Percentages of BrdU positive cells in the undifferentiated state ranged between 50-65% for *mock control, RET KO, ATP7A KO* and *SREBF1 KO* clones. In comparison to the *mock control* clone significantly less proliferative cells were detectable for the *ABCD1 KO* (~30% BrdU positive cells) (fig. 32 A‘).

In the differentiated state, no significant differences were observable comparing the respective edited clones to the *mock control* clone. Only for the *ABCD1 KO* a slight trend towards upregulated proliferative levels (adjusted p-value 0.08) was observable (fig. 32 A‘‘). The 7+1d differentiation state was selected for analyses to determine if cell clones differ in their ability to induce neuronal maturation. After 7d of cultivation in DM I (containing FBS), media was changed to DM II (contained no FBS). Corresponding to this, proliferative levels of clones were expected to be low after 7+1d in case neuronal cell fate induction was successful. Due to the experimental procedure of cell fixation, clones differentiated for longer periods could not be investigated accordingly as they tend to easily detach from the surface. Consequently, later differentiation stages were excluded from respective analyses.

Cell survival

The number of apoptotic cells in undifferentiated and 7+7d differentiated cells was determined by TUNEL assays (fig. 32 B). The percentage of $TUNEL^+$ cells in the undifferentiated state was below 1% for all clones except for the *ABCD1 KO* clone (~ 5%) with significant more $TUNEL^+$ cells in comparison to the *mock control* clone (fig. 32 B‘).

The amount of apoptotic cells within the differentiated cell populations was overall higher compared to the undifferentiated states, respectively. A significant increase in the $TUNEL^+$ cell fraction compared to the *mock control* clone was detected for the *SREBF1 KO* clone (fig. 32 B‘‘).

Cell migration

The cell migration capacity was assessed by a Boyden Chamber assay. Undifferentiated and 7+1d differentiated cells of all clones were investigated using a positive/chemoattractant and a negative/non-chemoattractant experimental setup, respectively (fig. 32 C). A significant reduction in the cell migration capacity of undifferentiated cells was detectable for the *RET KO* clone compared to *mock control* clone in the positive setup (using FBS as chemoattractant). All other investigated clones showed a similar migration capacity as the *mock control* clone (fig. 32 C‘).

Cell clones differentiated for 7+1d did not significantly differ in their migratory capacity. At this stage, BDNF was applied as chemoattractant in the positive control. However, high experimental variations indicated by larger SEM need to be considered. Mean numbers of migrated cells/membrane for differentiated cells were lower compared to undifferentiated numbers (fig. 32 C‘‘). Based on the experimental strategy, which required an enzymatic dissociation, later differentiation time points could not be evaluated accordingly.

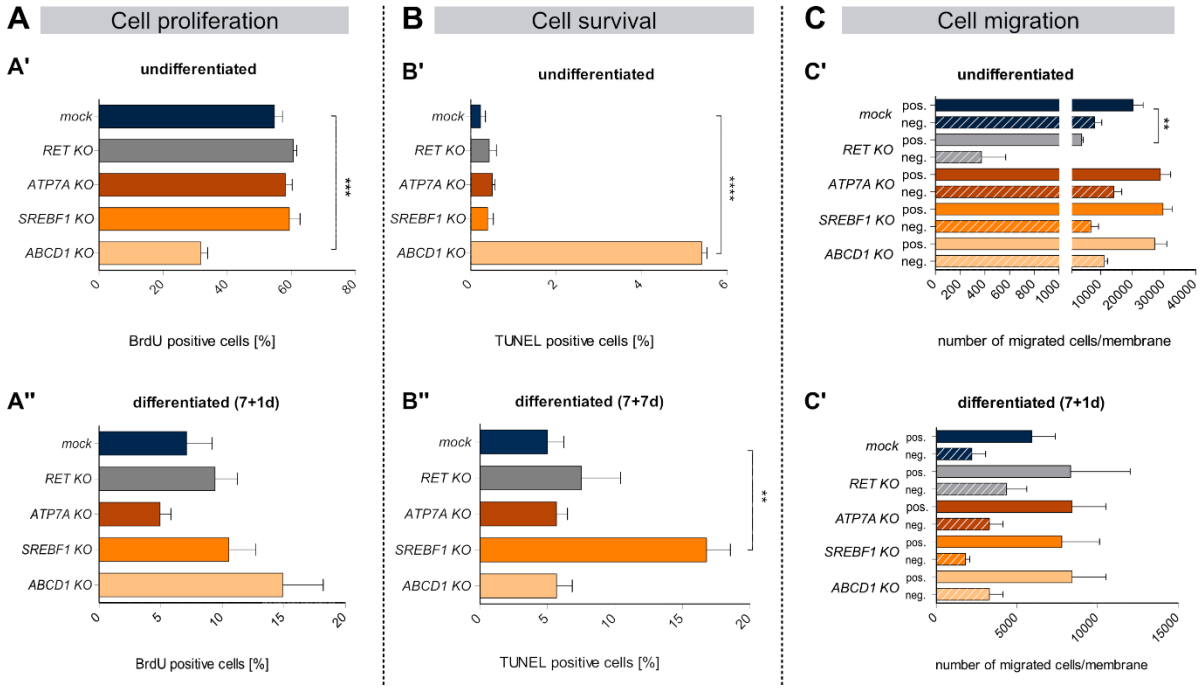


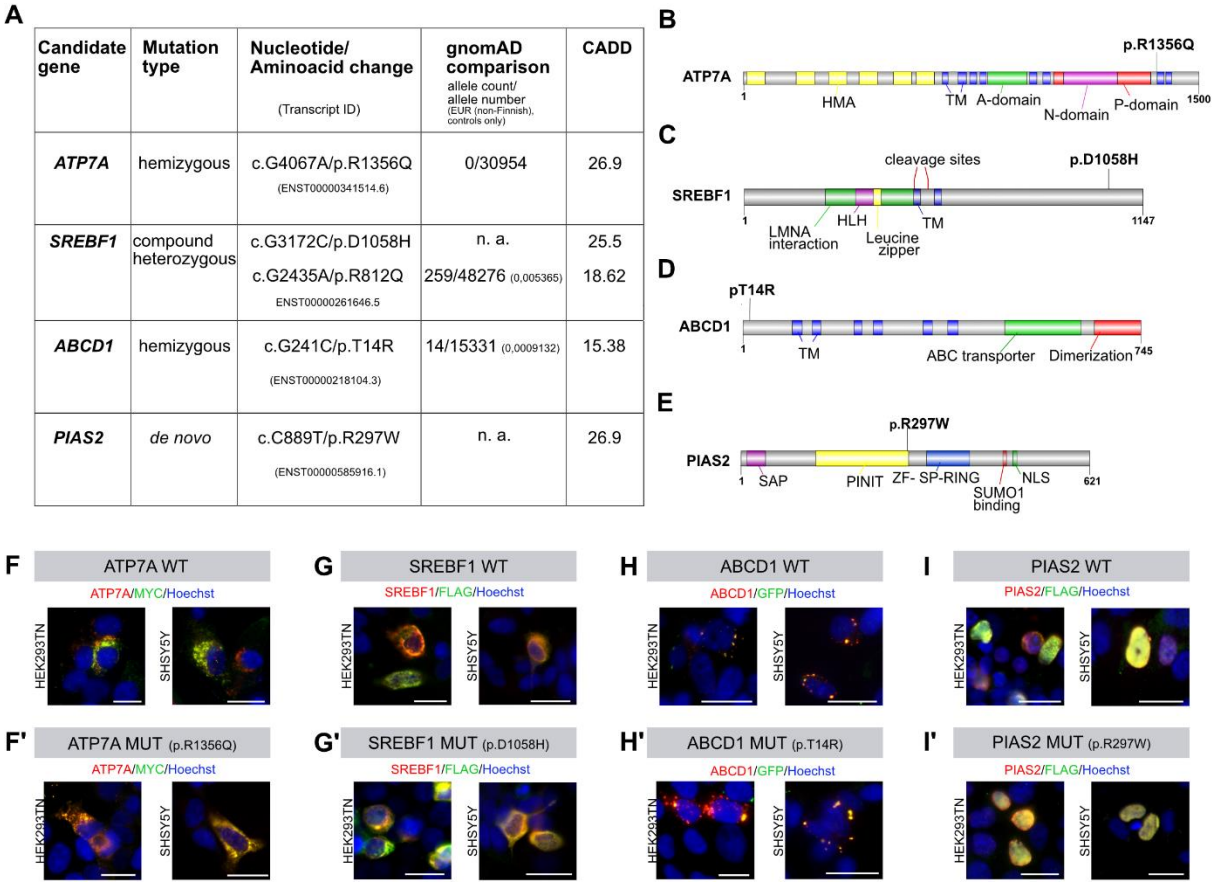
Figure 32: Comparative functional analyses of genome engineered clones compared to the mock control
A: The proliferation capacity of undifferentiated (A’) and differentiated (7+1d) (A’’) clones was assessed by BrdU Assays. **B:** The number of apoptotic cells in undifferentiated (B’) and differentiated (7+7d) (B’’) clones were determined by TUNEL assay. **C:** Cell migration capacity of undifferentiated (C’) and differentiated (7+1d) (C’’) cell clones was measured by Boyden Chamber assays. In the positive control, FBS was applied as chemoattractant for undifferentiated whereas BDNF was applied as chemoattractant for differentiated cells (negative: basal medium only). Bar plots show mean + standard error of mean (SEM) (n≥3), one-way ANOVA, Bonferroni corrected, *p<0.05; ** p<0.01; *** p< 0.001, **** p< 0.0001.

3.3.6 *Characterization of genetic variants*

Patient-specific mutations residing in selected candidate genes were analysed on different levels.
On genetic level, variant specific frequencies annotated in the gnomAD dataset (control individuals (non-Finnish Europeans only) were assessed. Except for the single nucleotide variant (SNV) in *ABCD1* (c.G241C) which was annotated in the database with a frequency of 0.0009132 none of the other variants was found in gnomAD. Moreover, the CADD score as a measure of deleteriousness of the genetic variant was considered. Respective scores ranged between 15.39 for the SNV in *ABCD1* and above 25 for the three other candidate variants (*ATP7A* (c.G4067A): 26.9; *SREBF1* (c.G3172C): 25.5; *PIAS2* (c.C889T): 26.9) (fig. 33 A).
On protein level, respective variants for domains for ATP7A, SREBF1 and ABCD1 were shown not to reside in any functionally annotated (fig. 33 B-D). Specifically, the variant detected in ATP7A (p.R1356Q) localised to the C-terminal part of the transporter protein adjacent to the second-last transmembrane

(TM) domain and in proximity to the phosphorylation domain (P-domain) (fig. 33 B). For SREBF1, the patient-specific missense variant localised to the C-terminal end of the protein (p.D1058H) (fig. 33 C), while for ABCD1 the amino acid substitution resided close to the N-terminus (p.T14R) (fig. 33 D). In contrast, the *de novo* variant in PIAS2 (p.R297W) caused an amino acid change in the most C-terminal end of the PINIT (Pro-Ile-Asn-Ile-Thr) region (fig. 33 E).

Subcellular localisation of recombinantly expressed WT and mutated (MUT) candidate protein versions were assessed by IF staining. In addition to the candidate protein, the specific tag staining revealed a co-labelling pattern. IF staining patterns of recombinantly expressed WT and MUT ATP7A indicated an endoplasmic reticulum (ER) or Golgi specific protein localisation. No differences between both protein versions were observable neither in HEK293TN nor in SHSY5Y cells (fig. 33 F and F‘).



Candidate gene	Mutation type	Nucleotide/ Aminoacid change (Transcript ID)	gnomAD comparison allele count/ allele number (EUR (non-Finnish), controls only)	CADD
ATP7A	hemizygous	c.G4067A/p.R1356Q (ENST00000341514.6)	0/30954	26.9
SREBF1	compound heterozygous	c.G3172C/p.D1058H	n. a.	25.5
		c.G2435A/p.R812Q (ENST00000261646.5)	259/48276 (0,005365)	18.62
ABCD1	hemizygous	c.G241C/p.T14R (ENST00000218104.3)	14/15331 (0,0009132)	15.38
PIAS2	*de novo*	c.C889T/p.R297W (ENST00000585916.1)	n. a.	26.9

Figure 33: Characterization of variants detected in the candidate genes on different molecular levels
A: Patient-specific variants in the selected candidate genes were assessed for their presence in the European (EUR) (non-Finnish) control population of the gnomAD dataset. The variant specific CADD score was calculated using the CADD model GRCh37-v1.4 (https://cadd.gs.washington.edu/snv). **B-E:** Protein domain structures of candidates and localisation of patient-specific mutations. Protein domains in specific isoforms were annotated based on database entries (Pfam, NCBI, Uniprot) or additional literature findings (Andreoletti *et al.*, 2017; Gourdon *et al.*, 2012) (ATP7A (B), SREBF1b (C), ABCD1 (D), PIAS2ß (E)). **F-I':** Subcellular investigation of wildtype (WT) and mutated (MUT) protein versions in HEK293TN and SHSY5Y cells transiently transfected with candidate specific overexpression constructs (ATP7A (F, F'), SREBF1 (G, G'), ABCD1 (H, H'), PIAS2 (I, I')). Candidate (red), protein-tag (green). Nuclei were counterstained with Hoechst (blue). Scale bar: 25µm. HMA (Heavy metal associated), TM (Transmembrane), A (Actuator), N (Nucleotide), P (Phosphorylation), HLH (Helix loop Helix), LMNA (LAMININ A/C), ABC (ATP binding cassette), SAP (scaffold attachment factor-A/B, acinus and PIAS), PINIT (Pro-Ile-Asn-Ile-Thr), ZF-SP-RING (Zinc Finger – Really Interesting New Gene), SUMO (small Ubiquitin-related modifier), NLS (Nuclear localisation signal)

Analyses for SREBF1 showed comparable results as described for ATP7A. Signal accumulation around the nuclei pointed towards a protein localisation within the ER. Staining patterns were comparable between WT and MUT proteins and between different cell lines, HEK293TN and SHSY5Y cells, respectively (fig. 33 G and G‘).
A vesicular-like staining pattern in HEK293TN and SHSY5Y cells was observable for recombinantly expressed ABCD1 WT. Introduction of the patient-specific variant into the overexpression construct did not cause any alteration in the subcellular localisation of ABCD1 (fig. 33 H and H‘).
The recombinantly expressed transcriptional co-regulator PIAS2 showed a nuclear restricted staining pattern in both investigated cell sources. Again, mutagenesis of the overexpression construct did not affect the protein localisation (fig. 33 I and I‘).

3.4 Candidate gene risk assessment

To evaluate the candidate gene-specific relevance for the HSCR aetiology, a scoring system was applied.
Database entries, literature findings and specific data generated within this project were summarized using a binary ranking system (tab. 9). Score Σ_1 was based on evidence generated in the categories: gene expression, information, interaction and constraint. In score Σ_2, additionally gene-specific mutation predictions were taken into account. For score Σ_3, significant differences detected in comparative functional *in vitro* analyses using the gene-specific *KO* cell clones of this project were considered (sections 3.3.4 and 3.3.5).

To allow a classification of the selected candidate genes of this project, risk scores for the established disease-causing factors *RET, SOX10, EDNRB, GDNF* and *PHOX2B* were calculated and implemented (data modified from Bachelor Thesis Dawid, 2019). Σ_2 scores for these known HSCR risk genes ranged between 8-13, with *PHOX2B* showing the lowest and *SOX10* displaying the highest risk score. *RET, EDNRB* and *GDNF* showed each a score of 12. However, no available information about *PHOX2B* in different databases must be considered. Additionally, no rare disease-causing variant was applicable and hence no scores could be attributed in the gene mutation category ($\Sigma_1=\Sigma_2$) (Bachelor Thesis Dawid, 2019).

ATP7A displayed the highest risk score among all selected candidate genes when only gene and mutation specific data was evaluated ($\Sigma_2=13$). *SREBF1* and *PIAS2* were annotated with a Σ_2 risk score of 11 and 12 each, while for *ABCD1* a Σ_2 score of 10 was calculated. Again, missing data in the Human Protein Atlas database must be considered for the evaluation of this candidate.
Taking *in vitro* function based data, respective Σ_3 risk scores for the candidates *ATP7A, SREBF1* and *ABCD1* as well as for the major HSCR susceptibility loci *RET* increased. Specifically, the candidate genes *ATP7A* and *SREBF1* as well as *RET* displayed a score Σ_3 of 14 each. For *ABCD1* a Σ_3 score of 13 was calculated. As no *KO* clone for *PIAS2* could be generated within this project and as *SOX10, EDNRB, GDNF* and *PHOX2B* were not included in the presented study pipeline, Σ_3 was not applicable for these genes.

Table 9: Final risk score evaluation of candidate genes in comparison to known HSCR-causing loci

RET, SOX10, EDNRB, GDNF and *PHOX2B* specific risk scores were modified from Bachelor Thesis Dawid, 2019. Details about gene-specific risk variants for *RET, SOX10, EDNRB* and *GDNF* which were used for risk scoring are annotated below. Patient-specific variants were evaluated for the candidate genes. - not annotated or not applicable. LoF (loss-of-function), oe (observed/expected), IBS/IBD (Irritable bowel syndrome/ Inflammatory bowel disease), maximal scores (Σ_1, Σ_2, Σ_3) are given in brackets

Category	**Gene expression**						**Gene information**			**Gene interaction**	**Gene constraint**		**Gene mutation**				**Functional *in vitro* analyses (3.3.5)**				
Classification/cut offs	mRNA level				Protein level		Function	Disease		Any association to ENS-related proteins	LoF oe value		CADD	Mutationtaster	PROVEAN		Cell differentiation	Cell proliferation	Cell survival	Cell migration	
	Human brain	Human colon	Human fetal colon	Murine embryonic ENS	Human Brain	Human Colon	Neuron specific	Neurological disease	IBS/IBD		<0.35	**Σ_1** (11)	>13	disease-causing	deleterious	**Σ_2** (14)					**Σ3** (18)
RET[1,*]	1	1	1	1	1	1	1	1	0	1	1	10	1	1	0	12	1	0	0	1	**14**
SOX10[2,*]	1	1	1	1	1	1	1	1	1	1	1	11	1	1	1	13					
EDNRB[3,*]	1	1	1	1	0	1	1	1	0	1	1	9	1	1	1	12					
GDNF[4,*]	1	1	1	1	1	0	1	1	1	1	1	10	1	1	0	12					
PHOX2B[*]	0	1	1	1	-	-	1	1	1	1	1	8	-	-	-	8					
ATP7A	1	1	1	1	1	1	1	1	0	1	1	10	1	1	1	13	1	0	0	0	**14**
SREBF1	1	1	1	1	1	0	1	1	0	1	1	9	1	1	-	11	1	1	1	0	**14**
ABCD1	1	1	1	1	-	-	1	1	0	1	1	8	1	1	0	10	1	1	1	0	**13**
PIAS2	1	1	1	1	1	0	1	1	0	1	1	9	1	1	1	12					

[1)]*RET:* Chr10:43597976 G>C [2)]*SOX10:* Chr22:38379456 C>T [3)]*EDNRB:* Chr13:78475316 C>A [4)]*GDNF:* Chr5:37816112 G>A
[1-4)] chromosomal positions refer to: Human GRCh37/hg19, *data modified from: Bachelor Thesis Dawid, 2019

4 DISCUSSION

4.1 Selection of candidate genes underlying HSCR

4.1.1 *Genetic heterogeneity on multiple levels*

HSCR has been classified as an oligogenic trait and lately evidence has been accumulated terming it a multifactorial condition. To date, more than 500 rare coding variants in the major HSCR risk loci (n > 20) have been identified (Tilghman *et al.*, 2019). However, the genetic etiology remains unclear for a large proportion of HSCR patients, as in fact in 70% of the cases the underlying genetics is yet unknown (Bahrami *et al.*, 2017). Moreover, authors of recently published sequencing studies claimed, that HSCR is most probably caused by a multitude of different rare and common, coding as well as non-coding variants residing in various genes that can act in concert (Luzón-Toro *et al.*, 2015b; Tilghman *et al.*, 2019; Zhang *et al.*, 2017).

Genetic profiles of two-sporadic L-HSCR patients had been assessed by WES and bioinformatic data filtering. This strategy indicated *ATP7A*, *SREBF1*, *ABCD1* and *PIAS2* as promising candidate genes (fig. 5). Large-scale reference datasets (ExAC and gnomAD browser) are now available and were used for bioinformatic data filtering. They are particularly useful, especially in the case of low-frequency variants (MAF), which were used to narrow down the most promising candidate genes. Consequently, these reference data can facilitate gene prioritization and selection of most promising causal variants in a disease context (Lek *et al.*, 2016). However, a general problem in genetic data filtering is the availability of control datasets. Ideally, respective controls should represent “healthy” individuals who are not affected by any phenotypic condition. Databases such as the ExAC and gnomAD include mostly genetic information from patients affected by various disorders ranging from metabolic syndromes to neuropsychiatric diseases (gnomAD browser; Lek *et al.*, 2016). Moreover, “non-diseased” individuals who served as controls in individual projects might present with other not reported phenotypes.

To find out if the two patients carry any rare variants in genes already described in the context of HSCR or ENS development, filtered WES data was compared to a dataset containing approximately 150 ENS/HSCR relevant loci (tab. S3, Supplementary). Patient II did not carry any rare variants in corresponding genes. In contrast, variants in *ECE1* (rs141146885), *SUFU* (rs34135067*)* and *SERPINI1* (rs61750375) were identified in patient I (fig. 5). *ECE1* is classified as HSCR risk gene (Luzón-Toro *et al.*, 2015b). This enzyme regulates maturation of the EDN ligands and therefore plays a crucial role in EDNRB signalling during ENS formation (Heanue and Pachnis, 2007; NCBI, Gene ID 1889). HSCR-associated variants in *ECE1* were so far only identified in a few patients (Hofstra *et al.*, 1999; Luzón-Toro *et al.*, 2015a). *SUFU* encodes a negative regulator of the SONIC HEDGEHOG (SHH) pathway, which is involved in the regulation of ENCDC proliferation, differentiation and migration during ENS formation (Fu *et al.*, 2004; Gui *et al.*, 2017; Nagy *et al.*, 2016; NCBI, Gene ID 51684). The synapse-associated serine proteinase inhibitor *Serpini1* has been described to be expressed in the developing murine ENS (Gui *et al.*, 2017; Heanue and Pachnis, 2006; NCBI, Gene ID 5274). While SNPs in *SUFU* and *SERPINI1* have so far not been reported in the context of HSCR, rs141146885 in *ECE1* has already

been detected in a male sporadic S-HSCR patient. This patient carried additional genetic variants in the HSCR susceptibility loci *RET* and *SOX10*. Furthermore, *in silico* analyses for the SNP in *ECE1* did not predict any functional relevance suggesting that this variant does not represent a major disease-causing factor (Luzón-Toro *et al.*, 2015a).

Except for a large number of coding variants, few common non-coding variants have been described in the context of HSCR (Emison *et al.*, 2005; Kapoor *et al.*, 2015; Tilghman *et al.*, 2019).
The SNP rs2435357 in intron 1 of the *RET* gene has been the first non-coding variant identified to significantly increase the HSCR risk. Specifically, the presence of the risk allele within the intronic enhancer region disrupts SOX10 binding and thereby reduces *RET* transactivation. This common low penetrance variant has been described to be present in about 80% of all HSCR cases (Emison *et al.*, 2010; Emison *et al.*, 2005). SNP genotyping showed that patient I carried the risk allele at rs2435357 in the heterozygous state while patient II revealed the risk variant in the homozygous state. Genotyping was also performed for two other *cis*-regulatory variants, rs7069590 and rs2506030, residing upstream of the *RET* transcriptional start site. Only for patient II, additional non-coding risk alleles were identified for rs2506030 while patient I carried the major, non-risk alleles at both *cis*-regulatory sites in a homozygous state (fig. 6 B). Risk alleles at both sites were shown to affect binding and target gene regulation by GATA BINDING PROTEIN 2 (GATA2) and RETINOIC ACID RECEPTOR BETA (RARB), respectively. Furthermore, a synergistic effect between all *RET* non-coding HSCR risk variants with regard to gene expression activation has been demonstrated (Chatterjee *et al.*, 2016). Due to its crucial role during ENS formation, reduced RET receptor expression contributes to the HSCR aetiology e.g. by impairing the migratory behaviour of pre-enteric NCCs at early stages or by reducing ENCDC proliferation at a later stage (Sasselli *et al.*, 2012).
The fourth genotyped non-coding SNP rs11766001resides in the *SEMA3* gene cluster. Although the functional impact of rs11766001 is so far not understood, this SNP has been classified as HSCR risk locus based on allelic association tests (Jiang *et al.*, 2015; Kapoor *et al.*, 2015; Tilghman *et al.*, 2019). Only patient I carried the risk allele in a heterozygous state at this locus (fig. 6 B).
Tilghman *et al.* reported recently that the disease risk was only significantly increased when at least five risk alleles in the four disease-associated non-coding SNPs (rs2435357, rs7069590, rs2506030, rs11766001) were present (Tilghman *et al.*, 2019). Taking the previously described data of both patients into account, none of them exceeded the proposed HSCR risk allele dosage for these common SNPs.

Besides genetic alterations on single nucleotide level, large structural variations have been rated relevant for HSCR (Jiang *et al.*, 2011b; Tang *et al.*, 2012). Different techniques are available to detect these large structural variations. For known candidate regions, CNVs can be identified e.g. by multiplex ligation-dependent probe amplification (MLPA) while comparative genomic hybridization (CGH) assays can be used as a hypothesis free approach. However, WES data is also suitable for CNV identification although it needs to be taken into account that only the coding-part of the genome is analysed (Fromer and Purcell, 2014; Pfundt *et al.*, 2017). CNV-based evidence for the two patients was obtained based on the generated WES data. After filtering, three rare CNVs in patient I and two rare CNVs in patient II were identified which were inherited in all cases (fig. 6 A).

Specifically, patient I carried a heterozygous deletion within the *PCDHG* gene cluster, which encodes for neural adhesion proteins probably involved in neuronal cell-cell connections (NCBI, Gene ID 56115). In this patient, 6 out of 22 genes, contained in that cluster, presented with the detected structural variation. *PCDHG* has so far not been linked to HSCR, but members of the alpha gene subfamily have already been shown to be expressed in the developing ENS and mutations in *PCDHA9* were just recently associated with the phenotype (Heanue and Pachnis, 2006; Shen *et al.*, 2018). Consequently, a functional relevance of *PCDHG* for HSCR is very likely. However, taking the mother's genotype into account, the presence of this structural variation is not sufficient to cause the disease. The patient carried additional copy number gains affecting *MHC CLASS I POLYPEPTIDE-RELATED SEQUENCE A (MICA)* and *UHRF1 BINDING PROTEIN 1 LIKE (UHRF1BP1L)* (fig. 6 A'). A direct association to HSCR has not been described for any of the genes but annotated tissue expression profiles and protein functions suggest a functional relevance for both in HSCR-causing mechanisms. The stress induced antigen *MICA* has been annotated to be expressed in the GI system and genetic variations have already been linked to CNS disorders such as Autism Spectrum Disorder (ASD) or schizophrenia (The Human Protein Atlas, MICA; Psychiatric Genomics Consortium, 2017; NCBI, Gene ID 100507436). For *UHRF1BP1L* an expression in brain and GI tissue has already been reported (The Human Protein Atlas, UHRF1BP1L). Taken together, SNVs at these loci might contribute to the individual's HSCR risk. However, before drawing any general conclusions, more detailed investigations are necessary.

For patient II, a copy number loss affecting the *CLN8 TRANSMEMBRANE ER AND ERGIC PROTEIN (CLN8)* and a copy number gain in the *GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE 3 (AGPAT9)* region were identified (fig. 6 A''). No HSCR associations have been reported, but their expression in brain and GI tissue annotated. Interestingly, *CLN8* underlies neuronal Ceroid lipofuscinosis, which is as well characterized by neurological symptoms as epilepsy (The Human Protein Atlas, CLN8; NCBI, Gene ID 2055; OMIM # 600143). Structural variants at both loci were inherited from the non-affected father, suggesting that their combined presence is not causing HSCR. However, further analyses are necessary as well as the recurrence of all detected CNVs affecting the genes needs to be assessed in additional HSCR patients.

In conclusion, CNV-based evidences were in line with published data, which suggest that not only syndromic but also non-syndromic HSCR subjects are frequently affected by structural variations (Tang *et al.*, 2012). Remarkably, some of the loci presenting with structural abnormalities have already been reported in the context of CNS-related disorders, such as ASD or neuronal Ceroid lipofuscinosis. In turn, ASD patients frequently present with GI symptoms (Rao, 2016). In this manner, commonalities in ENS and CNS regulatory molecular pathways and respective disorders are also reflected in the genetic data of this project (Tilghman *et al.*, 2019).

Combining all findings for the individuals of each family (fig. S1, Supplementary) and considering the previously outlined genetic heterogeneity in HSCR, leads to two conclusions. First, although parents in both families were not affected by the HSCR phenotype, their genotypes suggest the presence of GI dysfunctions. Deep phenotyping based on medical records and questionnaires to assess functional GI disorders using the Rome IV diagnostic questionnaire would be of particular interest (Palsson *et al.*, 2016). Second, albeit a variable number of probable HSCR risk loci could be identified in both patients

by CNV analyses and SNP genotyping, the presence of additional susceptibility variants in further, novel risk loci is very likely based on the multifactorial disease origin.

4.1.2 *Network analyses confirmed candidate selection*

To put the candidate genes into biological context and most recent knowledge, an *in silico* network analysis using the IPA knowledge base was performed. ATP7A, SREBF1, ABCD1 and PIAS2 were identified to be directly and indirectly linked to factors involved in ENS formation and/or HSCR aetiology (fig. 7).

An indirect connection was identified between ATP7A and RBP4 via the LDL complex, which has already been described in the context of atheriosclerosis (Qin *et al.*, 2010; Zhou *et al.*, 2018). RBP4 belongs to the lipocalin family where protein family members have been correlated with pro-atheriosclerotic processes (NCBI, Gene ID 5950; Oberoi *et al.*, 2015). Vitamin A (Retinol) has already been shown to be relevant in ENS development, as respective deficiency in *Rbp4*$^{-/-}$ mice causing distal bowel aganglionosis (Fu *et al.*, 2010).

IPA network analysis displayed a direct interaction between ABCD1 and DNMT3B, as this DNA methyltransferase has been identified to bind to CpG islands of the candidate gene's promoter (Jeong *et al.*, 2009). DNMT3B-mediated epigenetic regulation is known to be involved in NCC specification as well as in the onset of HSCR (Martins-Taylor *et al.*, 2012; Torroglosa *et al.*, 2014). A similar regulation of *ABCD1* might contribute to HSCR.

Direct interactions between NAV2 and PIAS2 or SREBF1 have been identified in different *in vitro* assays (Ebmeier and Taatjes, 2010; Rolland *et al.*, 2014). NAV2 plays a role in cellular growth and migration, and gene dosage effects have been reported in the context of HSCR (NCBI, Gene ID 89797; Turner, 2013). The transcriptional regulators SREBF1 and PIAS2 might be involved in the *NAV2* expression regulation. Furthermore, due to its function as E3 SUMOylation ligase, PIAS2 might affect respective protein stability by posttranslational protein modification (NCBI, Gene ID 6720 and 9063). This functional interaction applied as well to the direct interaction between PIAS2 and the ENS-relevant factor PTEN (Wang *et al.*, 2014). As the *Pten* gene deletion has been described to cause intestinal pseudoobstruction, posttranslational protein modifications mediated by PIAS2 might also contribute to HSCR (Puig *et al.*, 2009).

Additionally, SREBF1 was identified to be directly connected to the transcription factor *TBX3*, which is expressed in the developing ENS (Heanue and Pachnis, 2006; Lee *et al.*, 2014b).

In conclusion, the candidate gene selection made sense and could be integrated with known genes by IPA network analyses. The complex genetics and pathomechanisms of HSCR suggest, that candidates might represent modifiers affecting major ENS regulating genes and HSCR susceptibility pathways (Amiel *et al.*, 2008; Wallace and Anderson, 2011). Candidate specific protein functions will be addressed in more detail in the following paragraphs.

4.1.3 *Genetic based evidence for selected candidates*

To strengthen the putative relevance of the selected candidate genes on a genetic level, further rare variant carriers were identified for *SREBF1* and *PIAS2* in additional cases. However, the same variant as in the index patient could not be detected a second time (fig. 8). Nevertheless, CADD scores (> 13) suggested a functional relevance of all filtered, rare variants in the respective HSCR patients. In general, this score allows the unbiased annotation of various variants (SNVs as well as insertion/deletions) in the human genome by integrating data from existing tools in an innovative quantitative manner. Amongst others, annotations from conservation, population frequency as well as regulatory, functional and structural information are evaluated through a machine learning algorithm. A CADD score of 13 indicates that this variant is predicted to be among the 5% of most deleterious substitutions of the human genome (Kircher *et al.*, 2014).

ATP7A and *ABCD1* will be evaluated in the future according to the two other candidate genes. Interestingly, for both genes and for *PIAS2* rare variants were identified in patients suffering from chronic intestinal pseudoobstruction (personal communication Isabella Ceccherini, Genoa, Italy, data not shown). This fact points towards a relevance of these candidates in GI phenotypes.

To calculate if patients harbour a significantly increased number of rare variants in a specific candidate gene - compared to unaffected controls -, a genetic burden calculation could be performed in the future. However, genotype information used in respective studies of the International HSCR consortium for the control individuals were not accessible (section 2.2.1). The use of public data annotated in the ExAc or the gnomAD browser as controls is challenging, since e.g. sequencing platforms and variant-calling procedures differ between individual studies contributed to the browser data (Guo *et al.*, 2018). Multiple statistical methods are available to calculate the genetic burden, e.g. by testing for rare variant associations. However, these have different strengths and weaknesses with regard to robustness and statistical power and also require advanced bioinformatics tools as well as the availability of large datasets, which were not accessible in the context of this study (Lee *et al.*, 2014c).

4.1.4 *Expression patterns suggest a functional relevance of candidates in ENS formation*

mRNA expression validation of candidates in murine tissues

Due to its neurodevelopmental character, HSCR pathogenesis originates in early embryogenesis. Thus, assessing a candidate-specific relevance for the aetiology via expression analyses can only be achieved by using embryonic tissues. Transcriptional profiling interrogating more than 20 000 annotated genes in three murine, embryonic tissues was carried out. Specifically, specimens comprising pre-migratory vagal NCCs at stages E8.75 and E9.5 as well as the embryonic gut at stage E13.5 were analysed (fig. 9).

Vagal NCC delamination is initiated at E8.5 in mouse (Sasselli *et al.*, 2012). Therefore, dissected murine tissues at E8.75 and E9.5 include the respective pre-migratory vagal NCC population. This highly comparable cell type and tissue identity was displayed by means of hierarchical clustering of the overall mean expression values. As expected, correlation coefficients between samples of E8.75 and E9.5 were

higher than respective coefficients between E8.75 and E13.5 or E9.5 and E13.5 specimens, respectively (fig. 9 B).

Expression levels of the most prominent HSCR risk genes *Ret, Gdnf, Gfra1, Ednr, Edn3, Ece1, Sox10* and *Phox2b* confirmed tissue identity and reflected the molecular mechanisms underlying embryonic GI tract formation (fig. 9 C). Vagal NCCs at E8.75 and E9.5 mainly express *Sox10, Phox2b* and *Ret*, as these factors are responsible for the maintenance of multipotency (Sasselli *et al.*, 2012). Higher expression levels for these markers in early stages of ENS development (E8.75 and E9.5) compared to the latest investigated developmental time point (E13.5) indicate different facts. While increasing *Sox10* levels can be attributed to enteric glial cell fate commitment, upregulated expression values for *Phox2b* and *Ret* reflect enteric neuronal differentiation (Sasselli *et al.*, 2012).

The HSCR risk gene *Ednrb* is as well crucial for progenitor maintenance during early ENS development. However, in the literature an ENCDC-specific expression of *Ednrb* is described to be initiated only around E10.5 (Barlow *et al.*, 2003; Sasselli *et al.*, 2012). Increasing expression levels for *Ednrb* comparing tissues at E8.75, E9.5 and E13.5 might partially be reflected by this fact. Furthermore, this observation might as well indicate its crucial function during ENCDC migration, which is more pronounced at the latest investigated stage. The Ednrb-associated signalling factors *Edn3* and *Ece1* were also expressed in all investigated tissues. Again, increasing transcript levels might reflect their crucial roles as migratory guiding cues for Ednrb^{+} progenitors (Sasselli *et al.*, 2012). However, *Edn3* is not endogenously expressed in ENCDCs but gets secreted by the surrounding tissue, which might explain reduced gene expression values compared to *Ednrb* (Heanue and Pachnis, 2007). Comparable conclusions might be drawn for the *Ret*-specific ligand and the HSCR risk gene *Gdnf*. In contrast to Edn3, Gdnf does not only mediate rostro-caudal gut colonization by ENCDCs, but also gut invasion by vagal NCCs at E9.5 (Sasselli *et al.*, 2012; Young *et al.*, 2001). Interestingly, the *Ret* co-receptor *Gfra1* was expressed at higher levels than *Gdnf*, which most probably demonstrates that the first-mentioned factor is not only expressed by the surrounding gut mesenchyme, but also endogenously in ENCDCs (Chalazonitis *et al.*, 1998; Sasselli *et al.*, 2012).

The four selected candidate genes were expressed at comparable levels as the known HSCR risk genes, which points towards a putative relevance of these factors for HSCR and ENS development (fig. 9 C'). In the future, obtained transcriptional profiling data generated within this project will be useful as a valuable resource for HSCR candidate gene evaluation. However, these mRNA expression profiles are not suited to conclude about the functional relevance of a specific gene during ENS development. Additionally, it must be considered that dissected tissues did not only consist of ENS-specific cell types but contained various other cell lineages. Therefore, further analyses had to be carried out.

Protein expression analyses for detailed candidate characterization

Various protein expression analyses of the candidates were performed to complement transcriptional profiling results. Before jointly discussing all findings for each candidate, the purposes and limitations of the selected experimental strategies will be addressed in the subsequent paragraph.

Evaluation of the experimental strategies

IF analyses in the developing murine GI system were carried out to identify candidate-expressing cell types (fig. 10-13). Cell type markers selected for co-labelling with candidates illustrated proceeding GI tract formation and maturation.

At E9.5 and E10.5, multipotent pre-migratory vagal NCCs were identified by labelling with Sox10 and based on their position in close proximity to the neck. However, Sox10 is a general NC marker which is not specific to the vagal subpopulation (Simões-Costa and Bronner, 2015). At E11.5, $Sox10^+$ cells did not show any specific cell arrangement within the midgut and might either reflect gut invading, multipotent ENCDCs or immature glial cells (Lake and Heuckeroth, 2013; Sasselli *et al.*, 2012). As no additional immature glial marker was used for protein expression analyses, these lineages could not be differentiated in the respective analyses.

Three different pan-neuronal markers (Tubb3, Pgp9.5 and HuC/D) were used to detect the presence of immature enteric neurons at E11.5 and E13.5. While respective cells at the earlier stage were not connected to each other and did not display any cell organization pattern, a concentric neuronal cell layer indicative for the myenteric plexus was observable at E13.5. Formation of the submucosal plexus via a radial directed migration of ENCDCs did not occur before E15.5 and was therefore not observable in the respective investigations (Uesaka *et al.*, 2013).

Additionally, markers indicative for cell lineages residing in the ENS microenvironment were included into the candidate protein expression analyses. The importance of the ENS microenvironment has been investigated in chicken embryos. Specifically, ENS development was shown to be affected by inhibition of smooth muscle development and by Shh secretion from the epithelium (Graham *et al.*, 2017; Nagy *et al.*, 2016). The expression pattern of the smooth muscle marker Sma revealed the presence of the circular muscle layer in the developing midgut at E13.5. Smooth muscle differentiation is initiated when gut colonization by ENCDCs is still in progress, but in contrast to ENS development, generation of the circular and longitudinal muscle layer occurs by an outward-development (Fu *et al.*, 2004; Wallace and Burns, 2005). IF analyses of ECad illustrated a pseudostratified epithelium, which increased in thickness from E11.5 to E13.5 (Grosse *et al.*, 2011).

To obtain more experimental-based evidence for a putative neuronal specific role of the candidates, adult murine colon sections were also investigated for the expression of candidates. Enteric ganglia were visualized using the neuronal markers Tubb3 or HuC/D (fig. 14).

Although IF analyses offer the possibility to study protein expression patterns in detail, this kind of analysis has disadvantages as well. First, using tissue slices antigen co-expression in individual cells can only be assessed in case of an overlapping subcellular localisation of both proteins. Otherwise, a presumable co-expression can only be suggested based on IF signals in close spatial proximity to each other. Second, this technique requires a reliable antibody, which might not be available for all proteins of interest. Consequently, future studies focusing on the characterization of further novel HSCR candidate genes might apply multiplex RNAscope strategies instead or in an addition to IF analyses. This *in situ* RNA hybridization-based approach uses a specific probe design setup that allows signal amplification and simultaneous background suppression and therefore enables the detection of single molecules. Probes can be designed for any gene of interest and this technique can be easily applied at larger scales (Wang *et al.*, 2012).

Mechanisms and signalling molecules of ENS development are highly conserved between mouse and human, though there are also remarkable differences between both species. Comparative expression analyses in stage-matched human and murine tissues revealed differences in the expression patterns and dynamics of certain transcription factors (Memic *et al.*, 2018). Therefore, candidate protein expression patterns were as well investigated in human GI tissues in this study (fig. 16). Equivalent stage-matched human tissue sections to the analysed embryonic murine stages were so far not available but might be accessible in the future. Instead, fetal colon tissue was used, originating from a fetus acardius amorphous of the 25th week of gestation. At this developmental stage, the fetal gut has already reached its mature appearance with both plexus being present (Singh *et al.*, 2015). Nevertheless, due to its human specific prenatal background and as ENS maturation also continues postnatally, the generated data is highly valuable (Hao *et al.*, 2013; Rao and Gershon, 2018).
Quantitative assessment of candidates was performed by Western blot analyses using embryonic (E15.5) and adult GI and brain tissues. Findings were correlated with IF analyses in adult colonic ganglia to dissect a putative enteric neuronal specific role of the candidates (fig. 15).

Candidate protein expression patterns

Atp7a is a transmembrane protein, regulating intracellular copper transport as well as copper efflux in an Adenosine triphosphate (ATP)-dependent manner. This transporter is the main genetic determinant for Menkes disease – a multisystemic disorder characterized as well by progressive neurodegeneration (NCBI, Gene ID 538; Telianidis *et al.*, 2013).
IF analyses validated an Atp7a expression in or in close proximity to embryonic ENS-relevant cell types. These findings could be correlated with investigations in transgenic *Drosophila melanogaster* larvae as *DmATP7* has been expressed in neuronal and intestinal tissues – including the ENS (Burke *et al.*, 2008). Additionally, IF analyses suggested an Atp7a expression in enteric smooth muscle cells at E13.5. Protein expression in vascular smooth muscle cells has been reported previously; and on a functional level, a role in respective cell migration was suggested (Ashino *et al.*, 2010). This correlation might also apply to ENS formation.
Besides this cell type specific expression, *atp7a* has been shown to be ubiquitously expressed in early zebrafish embryos with most abundant signals in the notochord (Mendelsohn *et al.*, 2006). This data suggests a putative regulatory relevance of the copper transporter in early embryonic development.
Due to its relevance in Menkes disease, most detailed expression analyses for *Atp7a* have so far been performed in CNS-relevant tissues. Consequently, a multitude of different cell types e.g. neurons and glial cells, residing in various brain regions have been identified to express *Atp7a* (Telianidis *et al.*, 2013). An enteric neuron-specific expression in adult murine and human specimens was demonstrated in this study. However, ATP7A signals were also present in cells of the human intestinal epithelium. This was in line with its annotated copper delivery function to the circulation in adult enteric enterocytes (van den Berghe and Klomp, 2009). Western blot results generated in this project revealed decreasing Atp7a expression levels comparing embryonic and adult tissue specimens. Therefore, a predominant role of ATP7A in embryonic development and thus prenatal neurogenesis might be concluded.

Taken together, results of expression analyses further supported the selection of *ATP7A* as novel risk factor for HSCR by underscoring its putative regulatory function within the ENS microenvironment.

SREBF1 is classified as master regulator of cholesterol and fatty acid homeostasis and is synthesized as precursor (~120kDa) in the ER. Under low sterol levels, proteolytic cleavage releases the mature protein (~70kDa) that acts as bHLH transcription factor and regulates respective fatty acid and cholesterol signalling pathways (Eberle *et al.*, 2004; Horton *et al.*, 2002). Consequently, the observable nuclear restricted expression pattern in IF analyses indicated the presence of the mature Srebf1 protein. According to its relevance in lipid homeostasis, *Srebf1* transcripts have already been identified in liver, adipose and adrenal gland tissues (Im *et al.*, 2009; Shimomura *et al.*, 1997). The functional role of Srebf1 might as well explain its almost ubiquitous expression in embryonic ENS-relevant tissues. In fact, at these early developmental stages, the majority of cells - irrespective of their lineage – might be proliferatively active and hence have a high lipid demand. This hypothesis is in line with data in the respective literature, where SREBF1 expression was proven in proliferative tissues of the fetal intestinal crypt epithelium (Wilentz *et al.*, 2000). IHC staining in human tissues confirmed these results as prominent staining signals were observable within epithelial cells. In contrast to this, human enteric ganglia displayed only minimal protein expression. In adult murine colon tissue, only a subset of enteric neurons also co-expressed Srebf1. Combining results of IF and Western blot analyses, a role for Srebf1 in general neurogenesis might be suggested, since higher expression levels of Srebf1 were detectable in embryonic compared to adult tissues. A brain-specific expression of Srebf1 in rodents has already been reported (Ong *et al.*, 2000; Wilentz *et al.*, 2000). Furthermore, a functional relevance in this tissue can be suggested as *SREBF1* has already been reported in the context of different CNS related disorders like Parkinson's disease and schizophrenia (Chen *et al.*, 2017a; Do *et al.*, 2011).
In conclusion, generated protein expression data together with the association of *SREBF1* to CNS disorders, further strengthened the putative relevance of this candidate for the enteric neuropathy HSCR.

ABCD1 encodes for a VLCFA transporter, which regulates substrate uptake and its degradation by ß-oxidation in peroxisomes (NCBI, Gene ID 215). Dysfunctions of ABCD1 are responsible for the accumulation of VLCFAs and can result in X-linked adrenoleukodystrophy (X-ALD) – a neurodegenerative disorder (Baker *et al.*, 2015; Kemp *et al.*, 2011).
According to its function as peroxisomal transporter, IF analyses revealed a vesicular-like staining pattern. IF analyses proved an almost ubiquitous expression for Abcd1 in the investigated ENS-relevant tissues and therefore suggested a putative relevance for HSCR. Strachan *et al*. identified an *abcd1* expression in different progenitor cells as oligodendrocytes and motor neurons of early zebrafish embryos (Strachan *et al.*, 2017). An oligodendrocyte specific expression of Abcd1 in adult mouse brain tissue had already been described earlier. However, Troffer-Charlier *et al.* had demonstrated respective protein expression in a variety of different tissues and cell types (Troffer-Charlier *et al.*, 1998). More recent data has validated *ABCD1* expression in a broad range of different human tissues including the colon. Specifically, an expression in enteric ganglia of the myenteric and submucosal plexus as well as

in the intestinal epithelium has been presented (Höftberger *et al.*, 2007; Langmann *et al.*, 2003). The IHC analyses carried out in this project confirmed not only ABCD1-specific signals within adult colonic enteric ganglia, but also in intestinal epithelial cells. An enteric neuronal specific expression for Abcd1 was also validated in cryosectioned adult murine colon tissue. In Western blot analyses, GI specimens showed an enhanced *Abcd1* gene activity compared to CNS-containing tissues. However, deviating to mRNA data presented by Morvay *et al.*, Abcd1 protein expression levels were low in adult jejunum/ileum but higher in duodenum and colon samples. Abcd1 expression in the duodenum probably reflects its role in VLCFA turnover at this site (Morvay *et al.*, 2017; Mutch *et al.*, 2004). Most members of the ABC protein family function as dimeric transporters, which could also be illustrated in respective Western blot investigations, since two Abcd1-specific signals were detected (Morita and Imanaka, 2012).

Taken together, results of protein expression analyses as well as its known relevance for the neurodegenerative disorder X-ALD suggested a putative relevance of Abcd1 for ENS development and therefore the HSCR pathoaetiology.

PIAS2 acts as a SUMO E3 ligase for proteins of the SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) family as well as for various other transcription factors and proteins. Posttranslational SUMOylation can affect the target stability, interaction or activity (Kotaja *et al.*, 2002; NCBI, Gene ID 9063; Shuai, 2000). Pias2 is also described as transcriptional co-regulator and has already been shown to localise to nuclear speckles (Estruch *et al.*, 2016; Kotaja *et al.*, 2002). A definite nuclear-specific staining pattern was not observable in any of the IF or IHC investigations of this project. An explanation for this phenomenon remains to be found and explored in future studies. Nevertheless, there is increasing evidence for SUMOylation occurring outside of the nucleus - especially in the neuronal context (Martin *et al.*, 2007). Along the same lines, neuronal specific targets have already been described for PIAS2, such as the neurotransmitter nNOS (Shalizi *et al.*, 2007; Watanabe and Itoh, 2011). Pias2 revealed an almost ubiquitous expression in the investigated ENS tissues and thus a putative co-expression with all selected cell type markers. A widespread expression for *Pias2* at early embryonic stages was in line with published *in situ* hybridization data (Costa *et al.*, 2011). Moreover, the presence of *Pias2* in presumptive neuronal tissue has already been investigated in early *Xenopus laevis* embryos (Burn *et al.*, 2011). As demonstrated in this project, Pias2 expression was maintained as well in adult murine and human enteric neurons. Nevertheless, Western blot analyses suggested a predominant role for Pias2 during embryonic development, since respective specimens showed enhanced expression levels compared to the corresponding adult samples. In contrast to this, stable *Pias2* mRNA expression levels in the mouse cerebellum at different embryonic and postnatal stages have been described (Usui *et al.*, 2017). The CNS-specific expression of *Pias2* might also indicate the importance of the SUMOylation process in the context of neurological disorders. In fact, dysfunctions in this posttranslational protein modification pathway were already linked to Parkinson's, Alzheimer's or Huntington's disease (Anderson *et al.*, 2009; Rott *et al.*, 2017; Yang *et al.*, 2017).

In conclusion, the putative relevance of *PIAS2* for HSCR could be proven based on its expression in the ENS microenvironment and by taking its pivotal role in the SUMOylation process into account.

4.2 Detailed functional characterization of candidate genes

4.2.1 *Selection of a suitable cell culture model system – advantages and limitations*

As candidates were validated to be expressed at relevant stages and in relevant ENS cell types, their putative importance for the HSCR aetiology was further assessed by functional analyses in a neuronal cell model. Gene-specific *KO* cell clones of neuronal origin were generated using the CRISPR/Cas9 technology, and more specifically by gene disruption via the non-homologous end joining (NHEJ) mechanism. *KO* clones and a *mock* transfected and control-sorted clone were compared on various levels – with a special focus on cell features relevant in the HSCR pathogenesis such as differentiation, proliferation, cell survival and migration. In addition to the candidate genes, the major HSCR susceptibility locus *RET* was included as a *proof-of-principle* control.

Gene-specific *KO* cells were generated using the human neuroblastoma cell line SHSY5Y. This cell source does not resemble the ENS but allows a detailed gene characterization in a human, neuronal setup, which was the main purpose in this project. Furthermore, SHSY5Y cells are well accepted in the field as they were used to validating findings on a functional level in a number of HSCR studies (Sharan *et al.*, 2015; Shen *et al.*, 2018; Wu *et al.*, 2018). A human ENS-specific, commercially available cell line, which would mimic the *in vivo* situation at its best, is so far not available. The usage of human primary ENS cells would be of particular importance for future studies but first requires an intensive establishment and validation phase. In addition, applying the CRISPR/Cas9 genome editing technology to a primary cell source is much more challenging than in a cell line.
In summary, this project aimed to establish a pipeline for the functional neuronal *in vitro* characterization of selected HSCR candidate genes. Data might be used in the future as basis for advanced studies in a human primary ENS-derived cell culture system.

SHSY5Y cells were subcloned from the parental neuroblastoma cell line SK-N-SH which has initially been derived from a metastatic bone tumour biopsy of a four-year old girl (Biedler *et al.*, 1973). As neuroblastomas originate from embryonic cells of the NC, they share the same tissue origin as the ENS (Jiang *et al.*, 2011a; Johnsen *et al.*, 2019).
The SHSY5Y line is a heterogeneous cell population containing neuronal and epithelial-like cells which can be classified based on their morphology. Bidirectional interconversion between both cellular phenotypes has already been described (Ross *et al.*, 1983). This lineage heterogeneity is one of the major disadvantages that goes along in using this cell line for the specific purpose of this project. Individual gene-specific *KO* cell clones were classified on the basis of their neuronal-like morphology, while a cell clone classification based on the expression of the lineage marker *VIMENTIN*, which has been presented in different previous studies, was not successful (data not shown) (Bell *et al.*, 2013; Ciccarone *et al.*, 1989). Therefore, lineage-specific effects rather than KO-specific effects in individual genome edited clones cannot be ruled out. Analysis of further gene-specific *KO* clones might help to investigate and dissect this putative correlation.

4.2.2 *Gene-specific KO generation using the CRISPR/Cas9 technology*

All targeted genomic loci were endogenously expressed in the parental cell line SHSY5Y (fig. S4, Supplementary).

The *proof-of-principle* ***RET KO*** clone (c.231_232insC/ p.T78Hfs5*) was validated for the respective protein loss in the context of this project (fig. 17 A). So far, no *RET*-specific *KO* clone was described, but RNAi knockdown analyses for this gene have already been performed in the SHSY5Y cell line (Esposito *et al.*, 2008; Meka *et al.*, 2015).

Three out of four candidate genes were successfully knocked-out in the context of this project. The ***ATP7A KO*** (c. 500_501insT/500_501insA; p.Q167Hfs*4/p.A168Sfs*3) clone, the ***SREBF1 KO*** (c.874_875insT/c.874_875insTT; p.K293Efs*70/K293*) clone and the ***ABCD1 KO*** (c.1004_1005insG; p.K336Efs*65) clone either harboured compound heterozygous or homozygous genome modifications at the targeted loci (fig. 17 B-D). The respective protein loss was validated by Western blot analyses. However, as in the case of SREBF1, low endogenous expression levels, many annotated protein coding isoforms and a low-quality antibody impeded this validation step. Therefore, protein lysates of HEK293TN cells transiently transfected with a gene-specific overexpression construct served as internal controls in all immunoblot investigations. *KO* specific clones for *ATP7A, SREBF1* and *ABCD1* in the SHSY5Y cell line have not been described in literature to date.

CRISPR/Cas9 mediated gene disruption through the introduction of a premature stop codon might cause the generation of a truncated protein version of unknown functional relevance (Reber *et al.*, 2018). This applies only to edited transcripts that are not targeted by nonsense-mediated decay (NMD), which was not studied in the context of this project. However, the presence of truncated protein version is unlikely as no additional bands were observable in respective *KO* clone protein lysates. Nevertheless, only mass spectrometry analyses are suitable to rule out their presence (Zhang, 2019).

KO clones were assessed for off-target effects because high off-target frequencies of about 50% were reported in the context of RNA-mediated Cas9 guiding (Zhang *et al.*, 2015). Ideally, off-target detection in *KO* clones should be performed by next-generation sequencing. However, due to its high costs, only selected off-targets were assessed by Sanger sequencing in this study. Specifically, all exonic regions predicted in the top 20 off-target sites by the *CCTop CRISPR/Cas9 target online predictor* were included into sequence validation (Stemmer *et al.*, 2015). Furthermore, predicted intronic and intergenic regions were assessed for their putative functional relevance. In this regard, a high sequence conservation across different species and/or the presence of active chromatin marks, annotated expressed sequence tags (ESTs) or DNase clusters were valid criteria for a selection for Sanger sequence validation.

Off-target effects were excluded in all gene-specific *KO* clones, except for the *SREBF1 KO.* For this clone, a heterozygous one-nucleotide insertion in intron 7 of the *EEF1A2* gene was identified (fig. S7, Supplementary). However, detailed *in silico* analyses excluded a putative functional relevance as the annotated mRNA EST was located more 3‘ than the introduced nucleotide, and the annotated DNase cluster was so far not reported in the SHSY5Y cell line (data not shown). Consequently, the *SREBF1 KO* was finally validated for its further usage.

In general, this data demonstrated the necessity for detailed investigation and evaluation of potential off-target effects when applying the CRISPR/Cas9 methodology.

Genome editing for the candidate gene ***PIAS2*** was not successful as no gene-specific *KO* clone could be generated. In a first experimental setup, exon 6 of *PIAS2* has been targeted and led to the generation of an edited clone as demonstrated by genotyping (c.774_775insA) (Master Thesis Mederer, 2015). Detailed investigations on transcript level revealed alternative splicing of the edited exon (fig. 18 A). As exon 6 consisted of a multiple of 3 bases, alternative splicing did not disrupt the open-reading frame of the gene and a novel protein coding isoform might have been generated. This was not followed up in this project, but database research revealed that a multitude of protein coding isoforms had already been annotated for this gene (tab. S4, Supplementary).

The phenomenon of alternative splicing in the context of CRISPR/Cas9 mediated genome editing has been reported in several recent studies (Chen *et al.*, 2018; Lalonde *et al.*, 2017; Mou *et al.*, 2017; Uddin *et al.*, 2015). Nevertheless, precise knowledge of the molecular causes triggering this mechanism is still missing (Smith *et al.*, 2018). Based on the *PIAS2*-specific finding, the general editing strategy of this project was revised and all other gene-specific *KO* clones were validated for no alternative splicing of the edited exon (fig. S6, Supplementary).

In a second editing setup, sgRNAs targeted different sites in exon 2 of *PIAS2,* which did not consist of a multiple of 3 bases (fig. 18 B) (Addgene, 2017; Lalonde *et al.*, 2017). However, still no *KO* clone could be generated. Specifically, the lineage heterogeneity of the parental cell line SHSY5Y caused problems as almost exclusively clones of epithelial-like morphology survived the single cell cloning procedure. The putatively edited clone of neuronal-like morphology, generated by using sgRNA III, carried most probably a larger genomic rearrangement at the Cas9 target site. This would explain the failure of generating an amplicon for DNA genotyping and missing transcript expression of the *FL PIAS2ß*. The unintended generation of large deletions and chromosomal rearrangements by CRISPR/Cas9 genome editing was just recently explored (Kosicki *et al.*, 2018).

Gene *KO* escape when using the CRISPR/Cas9 technology was described for a mitotic spindle checkpoint kinase. Here, the authors claimed that this phenomenon might be explained by a near essential cellular role of the targeted locus (Meraldi, 2019). CRISPR/Cas9 editing of *PIAS2* in another cell line might help to elucidate if this hypothesis applies as well for the selected candidate gene. However, if a near cell essential or neuronal specific role could be proven experimentally, a functional relevance for *PIAS2* in HSCR would be very likely.

4.2.3 *Impairments of KO clones in cell features relevant for ENS development*

Validation of neuronal-like *in vitro* differentiation of SHSY5Y cells

The *KO* clones of *RET*, *ATP7A*, *SREBF1* and *ABCD1* were validated for their suitability for a more detailed functional characterization. To assess their neuronal differentiation capacity, a suitable differentiation protocol was required, which was established based on the parental cell line SHSY5Y.

Neuronal cell fate induction in the heterogeneous cell population SHSY5Y causes the generation a more homogenous neuronal-like phenotype (Encinas *et al.*, 2000). Two pilot experiments (A, B) were carried out, using modified versions of published differentiation protocols (fig. 19 and 20). Specifically, three different coatings and two different media compositions were tested as the influence of both for neuronal *in vitro* differentiation has already been characterized (Agholme *et al.*, 2010; Encinas *et al.*, 2000;

Raghavan and Bitar, 2014). As neuronal differentiation induction by the Vitamin A derivative RA is a well-established method for SHSY5Y cells, this treatment was included in all experimental setups (Kovalevich and Langford, 2013). In general, a differentiation protocol should be established for different functional *in vitro* assays in the gene-specific *KO* clones. The neuronal subtype, that was predominantly generated, was not of a specific interest for this project as in the HSCR phenotype all subtypes of enteric neurons are completely missing.

The final selection of differentiation **protocol Ib** for all further analyses was mostly based on morphological evaluations by brightfield microscopy, as mRNA expression profiles displayed only minor differences comparing all tested protocols.

mRNA expression results for WT SHSY5Y cells differentiated with protocol Ib confirmed the suitability of the selected protocol. Specifically, relative transcript levels for the neuronal progenitor markers *P75NTR* and *NES* mostly decreased – especially during early phases of differentiation. A downregulation of the intermediate filament protein NES in the context of neuronal differentiation has already been reported (Murillo *et al.*, 2017). It needs to be determined in further analyses if upregulated relative expression levels of *P75NTR* after 7+14d of differentiation reflect its role in glial cell commitment or its relevance in apoptosis regulation e.g. by expression analyses of additional cell fate markers (Meeker and Williams, 2015). As expected, many neuronal cell fate markers, such as *SYP, MAP2* and *GAP43,* showed increasing transcript levels during differentiation. These observations were in line with literature findings, since an upregulation on protein level has already been analysed for MAP2 and GAP43 in differentiating WT SHSY5Y cells (Encinas *et al.*, 2000; Murillo *et al.*, 2017). The pan-neuronal markers *UCHL1* and *TUBB3* were found to be almost stably expressed during differentiation, which also partially confirmed previous protein-based investigations (Constantinescu *et al.*, 2007; Murillo *et al.*, 2017). Interestingly, transcriptional profiling of *ASCL1* demonstrated a strong reduction in relative mRNA levels comparing undifferentiated and early differentiated SHSY5Y cells. This observation coincided with published data from Lopez-Carballo *et al.*. Authors suggested that a change in the balance between differentiation promoting and inhibiting factors might be crucial for general neuronal lineage induction in this cell source (Lopez-Carballo *et al.*, 2002). Similar observations were made for *NPY*, although a high standard deviation of the undifferentiated sample needs to be considered. Nevertheless, this neuronal subtype specific marker was as well found to be upregulated during later neuronal cell fate induction and thereby confirmed previous data (Goldie *et al.*, 2014). In contrast to *NPY*, the second neuronal subtype specific marker *TAC1* was found to be almost unexpressed. This applied as well to the progenitor marker *SOX10* and has already been reported previously (Acosta *et al.*, 2009). Relative mRNA expression levels of the microtubule associated protein *TAU* did not reflect proceeding neuronal maturation but mRNA expression levels for this marker have been shown to only poorly correlate with respective protein levels and thus neuronal maturation stages (Agholme *et al.*, 2010). This fact might as well apply to the results in this project.

On morphological level, differentiating WT SHSY5Y displayed proceeding neuronal-like network formation, since progressive neurite extension and an increase in the number of neuronal-like branches were observable. These morphological features are generally regarded as valuable properties indicating neuronal-like maturation in SHSY5Y cells (Encinas *et al.*, 2000; Kovalevich and Langford, 2013).

***In vitro* characterization reveals functional defects of variable degrees in the *KO* clones**

Differentiation of all cell clones using the previously validated protocol, was limited to 7+28d. Although transcriptional profiles of the differentiating WT SHSY5Y cell populations suggested that no complete neuronal maturation was reached at this time point, this stage was considered to be a suitable compromise between time and neuronal-like maturation degrees.

The set of markers used for mRNA expression profiling in the cell clones was further expanded, as the apoptotic marker *CASP3*, the candidate genes and the major HSCR susceptibility locus *RET* were included. To assess if the marker expression repertoire might be altered in lineage-specific SHSY5Y clones, the glial marker *GFAP*, which has so far not been reported in the context of WT SHSY5Y cell differentiation, was also implemented (Kovalevich and Langford, 2013). *GFAP* expression analyses might be also used to classify a putative glial cell specific activity of *P75NTR* at later stages of differentiation (Cragnolini and Friedman, 2008; Meeker and Williams, 2015). In addition to expression analyses on mRNA level, markers indicative for proceeding neuronal-like maturation and neuritogenesis were selected for IF staining.

Results for expression analyses on mRNA and protein level as well as brightfield microscopy confirmed the suitability of the chosen differentiation protocol for neuronal differentiation of the lineage-specific ***mock control*** clone (fig. 21 and 26).

On mRNA level, only minor differences were detectable between the *mock control* clone and the parental cell line. *SOX10* and *GFAP* were also not expressed in the lineage-specific *mock control* clone. Additionally, relative transcript levels for the pan-neuronal markers *TUBB3* and *UCHL1* in the control clone were as well almost stable throughout differentiation. Minor differences comparing both cell identities were observable for the two neuronal progenitor markers *P75NTR* and *NES*. Specifically, expression levels for the neuroepithelial-specific stem and progenitor marker *NES* did not decline during differentiation, and for *P75NTR* overall higher relative transcript levels were rather observable in the *mock control* clone than in the parental line. These discrepancies might most probably be attributed to the difference in cell lineage composition of both entities. It needs to be determined if higher *P75NTR* relative expression levels thereby reflect the presence of more neuronal-like characteristics and an enhanced neuronal-like maturation capacity in the clone.

Proceeding neuronal *in vitro* maturation of the *mock control* clone was displayed on various levels. First, the relative *P75NTR* mRNA expression level decreased while relative transcript levels of almost all neuronal markers (*UCHL1*, *TUBB3*, *TAU*, *MAP2*, *SYP*, *NPY)* increased under differentiating conditions. Second, brightfield microscopy analyses illustrated a gradual increase in the length of neurite-like outgrowths and a gain in the number of neuronal-like branches and synaptic-like contacts under differentiating conditions. Extensive cellular clustering, which was observable after 7+14d, was in line with previous literature findings (Constantinescu *et al.*, 2007). After 7+28d, this clone formed a fragile network of clusters connected by long, semi-adherent processes that resembled axons. Third, protein expression analyses of selected neuronal markers confirmed mRNA expression data and conclusions drawn from brightfield microscopy. Corresponding to their function as cytoskeleton proteins, the neuronal markers TUBB3, TAU, MAP2 and GAP43 localised to neurite-like outgrowths. As expected, the synaptic marker SYP generated a dotted-like staining pattern along these neurites. Additionally, only

a minority of differentiated cells were positive for the proliferative marker Ki67. This suggested that neuronal-like clusters rather arose from migrated cells, but were not generated by extensive proliferation.
The expression profile of *RET* in the *mock control* clone confirmed previous literature findings, as RA treatment has been demonstrated to cause an upregulation in respective gene expression, which might be crucial for general neuronal cell fate induction in the SHSY5Y cell line (Bunone *et al.*, 1995; Esposito *et al.*, 2008; Korecka *et al.*, 2013). Interestingly, all candidate genes displayed a similar expression profile as *RET*, which might point towards a comparable regulatory role of these genes in neuronal maturation.

All *KO* clones were differentiated and analysed according to the *mock control* clone. Comparative functional *in vitro* assays targeting the cell proliferation, survival and migration rate were additionally carried out (fig. 32). Results of all these investigations will be jointly discussed for each of the gene-specific *KO* clones in the subsequent paragraphs.

The ***RET KO*** represented the *proof-of-principle* control in this study, since *RET* is the major HSCR susceptibility locus involved in the regulation of all important ENCDC functions (differentiation, proliferation, cell survival and migration) during ENS formation (Heanue and Pachnis, 2007).
Detailed functional characterization of the *RET KO* clone displayed most striking phenotypic alterations for this clone in its neuronal differentiation capacity and its migration capability (fig. 22, 27 and 32). This is in line with *in vitro* data of HSCR patient iPSC-derived ENCDCs (carrying a *RET* mutation), which showed most prominent defects in migration and neuronal lineage commitment (Lai *et al.*, 2017). Moreover, neuronal precursors of *Ret*-deficient mice have shown impairments in migration and neuronal differentiation too (Enomoto *et al.*, 2001).
Neuronal network formation in the *RET KO* was delayed in comparison to the *mock control* clone, as neuronal-like clusters were not arranged in a highly organized network before 7+21d of differentiation. On the expression level, a striking difference comparing the *RET* with the *mock control* clone was detectable for the neuronal progenitor marker *P75NTR*. Specifically, RA treatment during the first week of *in vitro* differentiation caused a strong upregulation for this marker in the *RET KO* clone, which was not observable for the *mock control* clone. A functional interaction for Ret and p75 was already described in the context of cell-death regulation in sympathetic neurons and with regard to sensory neuron diversity generation (Chen *et al.*, 2017b; Donnelly *et al.*, 2018). If RET deficiency in the *KO* clone is mechanistically and functionally connected to this difference in the *P75NTR* expression profile, needs to be investigated.
As revealed by descriptive statistical analyses, relative *UCHL1* mRNA expression levels were detected to be significantly reduced and additionally showed a gradual increase during differentiation in the *RET KO* clone. This observation might reflect the delay in neuronal-like maturation, as described previously. Further neuronal markers were as well found to be differentially expressed in comparison to the *mock control* clone. For *MAP2* and *GAP43* differences were observable at various stages of neuronal differentiation. The last-mentioned gene was already shown to be downregulated in the context of a transient *RET* knockdown in SHSY5Y cells, which manifests the failure in neuronal differentiation in respective cells (Esposito *et al.*, 2008).

As downstream signalling pathways of *RET* are highly diverse, further investigations are required to uncover underlying mechanisms and functional links that underlie the differential marker expression in this *KO* clone (Ibáñez, 2013; Kurokawa *et al.*, 2003). However, qRT PCR data was so far only analysed on an exploratory level in order to obtain a first impression about putatively misregulated marker genes. Therefore, no confirmatory conclusions can be drawn. Additionally, observations on mRNA level should be validated as well on protein level. Although respective IF staining revealed expression and correct subcellular localisation of the investigated neuronal markers, these analyses are not suitable for any quantitative conclusion. Instead, Western blot analyses might be carried out in the future.

Interestingly, except for *ATP7A* in the 7d samples, none of the candidate genes was differentially expressed in the *RET KO* clone compared to the *mock control* clone. This observation suggested that candidate genes are probably no downstream targets of the RET signalling pathway but might either act upstream or might be involved in other ENS-relevant signalling cascades.

Besides the alterations in the neurogenesis behaviour, the *RET KO* clone showed also a failure in migration in its undifferentiated state compared to respective *mock control* cells. The relevance of *RET* for the migratory capacity in SHSY5Y cells has already been described (Cockburn *et al.*, 2010). However, this phenotypic impairment was not reproducible in the differentiated state of the *RET KO* clone. With regard to this, two facts need to be considered. First, RA treatment applied in the context of neuronal *in vitro* differentiation decreases the general migration capacity of neuroblastoma cells (Messi *et al.*, 2008). This is in line with results of Boyden Chamber assays, since mean numbers of migrated cells/membrane were lower for differentiated than for undifferentiated cell clones. Second, high inter-experimental variations displayed by large SEMs might mask putative differences between the individual clones. This variability was most probably caused by the experimental procedure as an enzymatic dissociation of 7d neuronal-like cells was necessary; but these cells were highly sensitive and delicate to this procedure due to their advanced maturation state.

In conclusion, *RET KO* cells displayed defects in two HSCR-relevant cell features - differentiation and migration - thereby partially reflecting the disease-causing pathomechanisms in HSCR patients. Based on these findings, the neuronal-like cell culture system was evaluated as suitable model to also assess novel HSCR candidate genes for their relevance in the pathoaetiology.

For the ***ATP7A KO*** clone most striking alterations were detected in its *in vitro* differentiation behaviour based on microscopic analyses and mRNA expression profiling (fig. 23 and 28).

On morphological level, at any time of neuronal differentiation, *ATP7A KO* cells were arranged in a neuronal-like network comparable to that of the *mock control* clone. In fact, cell densities for this clone were considerably higher than for any other clone throughout differentiation. Additionally, many cells were apoptotic and floating, which caused problems in the cell fixation after 7+28d of differentiation. The protocol for IF analyses might be further modified for this clone in the future to prevent cell detachment.

Induction of neuronal maturation should be associated with a downregulation in proliferation – especially when differentiating cells are cultivated without any serum (SHSY5Y DM II). ATP7A expression is inversely correlated with neuroblastoma cell proliferation (Telianidis *et al.*, 2013). This functional interconnection might account for higher cell densities observable in the respective *KO* clone. IF

analyses after 7+28d validated this hypothesis, since many *ATP7A KO* cells were immunoreactive for the proliferative marker Ki67. However, proliferation rates assessed by BrdU assays did not significantly differ from observations of the *mock control* clone – neither in the undifferentiated nor in the differentiated state. Analyses at further differentiation stages should be carried out to investigate if comparable results can be generated. Moreover, it needs to be stated that Ki67 expression is highly dynamic during the cell cycle (Miller *et al.*, 2018). Higher cell densities might as well be caused by reduced cell death although TUNEL assays did not reveal any significant alterations in this clone in comparison to the control. However, mRNA expression levels for the apoptotic marker *CASP3* in the *ATP7A KO* clone were almost stable throughout differentiation and - as suggested by exploratory data analysis – transcript values were significantly lower compared to the *mock control* clone. Nonetheless, the protease CASP3 is not only associated to regulated cell death but might as well play a role in neurogenesis and synaptic activity (D'Amelio *et al.*, 2010). To overcome this issue in prospective experiments, expression analyses might be performed for members of the pro-apoptotic *B-CELL LYMPHOMA 2 (BCL2)* gene family, e.g. *BCL2-ASSOCIATED X (BAX)* (Westphal *et al.*, 2011).

Microscopic analyses suggested an impaired neuronal-like maturation capacity of the *ATP7A KO* clone. This *in vitro* phenotype might mimic the neurodegenerative disorder Menkes disease. Protein deficiency in the disease context results in defective cellular copper transport and thus in functional impairments of various - as well as neuronal-specific - cuproenzymes (Kaler, 2011; Telianidis *et al.*, 2013). In *Atp7a* mutant mouse lines, impairments in proper axonal outgrowth and synaptogenesis have been described (El Meskini *et al.*, 2007). Based on the exploratory mRNA expression data analysis, the *ATP7A KO* clone revealed significant downregulations in the expression of neuronal cell fate markers, such as *MAP2* and *GAP43*, at multiple time points of differentiation. Corresponding to murine KO data, these differences might also reflect a functional impairment in proper axonal outgrowth for the *ATP7A KO* clone. Moreover, *TAU* expression levels were higher in the undifferentiated and 7d differentiated samples than at later differentiation time points. These findings on mRNA level were confirmed on protein level as only weak and disorganized TAU-specific IF signals were detectable in differentiated cells.

Although functional *in vitro* characterization of the *ATP7A KO* clone revealed impairments in the neuronal-like differentiation behaviour, further investigations are necessary to draw more detailed conclusions and to clearly identify the functional relevance of *ATP7A* for proper neurogenesis. Specifically, further mRNA expression analyses applying advanced statistical data evaluations as well as additional quantitative protein expression investigations need to be performed.

For the ***SREBF1 KO*** clone most prominent alterations were detectable with regard to the overall differentiation behaviour and the cell survival rate after 7+7d of differentiation (fig. 24, 29 and 32). Comparable to the *proof-of-principle* control, neuronal-like network formation was delayed. Additionally, the *RET KO* and the *SREBF1 KO* clones shared similarities in their mRNA expression profiles, e.g. for the neuronal progenitor marker *P75NTR* and the pan-neuronal marker *UCHL1*. These findings might point towards a functional relevance of SREBF1 during early neuronal maturation and general neurogenesis in SHSY5Y cells, as it was previously suggested for RET.

As outlined in 4.1.4, *SREBF1* has already been associated to different neurological conditions which are partially characterized by myelin abnormalities. Its importance for autonomous lipid synthesis within the nervous system has already been highlighted in different studies. It has been shown that *Srebp1c* KO mice showed a peripheral neuropathy and that blocking of this factor caused defects in oligodendrocyte maturation (Camargo *et al.*, 2009; Cermenati *et al.*, 2015; Monnerie *et al.*, 2017). In *Drosophila* larvae, srebp has been identified as a master regulator for neuronal lipogenesis, reflected by its relevance for proper dendrite formation (Ziegler *et al.*, 2017). As these studies point towards a pivotal role of Srebf1 in neuronal differentiation and neuronal network formation, the observable impairments of the *SREBF1 KO* clone in neuronal-like *in vitro* differentiation might be interpreted accordingly.

This clone showed a significantly increased apoptotic rate after 7+7d of differentiation in comparison to the *mock control* clone. In literature, SREBF1 has already been identified to be crucial for general cell growth and cell viability in cancer cells (Griffiths *et al.*, 2013; Porstmann *et al.*, 2008).

Combining both regulatory functions of SREBF1 might explain the decreased cell viability of this clone under differentiating conditions. SREBF1 deficiency might result in defective lipid metabolism, which prevents neurite elongation and hence results in enhanced apoptosis of 7+7d differentiated cells. Interestingly, exploratory mRNA expression data analysis suggested significantly reduced mRNA levels of *MAP2* and *GAP43* – two factors associated with neurite outgrowth - at the same stage in the *SREBF1 KO* clone. The proposed hypothesis would be in line with results of brightfield microscopy analyses, since *SREBF1 KO* cells were not connected by extensive long and semi-adherent neurite-like after 7+14d.

Further investigations are necessary to determine how the loss of SREBF1 contributes in detail to the observable phenotypic alterations. To confirm the identified deregulated cell type markers, a targeted experimental setup evaluated with advanced statistical methods should be applied. Additionally, expression changes on mRNA level need to be verified on protein level. IF analyses for all investigated markers proved respective protein expression and confirmed their expected subcellular localisation but no quantitative conclusions can be drawn from these investigations.

Functional *in vitro* characterization of the ***ABCD1 KO*** clone demonstrated only minor differences in neuronal differentiation behaviour in comparison to the *mock control* clone, while most striking phenotypic alterations were detected for the proliferation and cell survival rates (fig. 25, 30, 32). Specifically, in the undifferentiated state the proliferative rate was significantly reduced, and the apoptotic rate showed a significant increase for this clone.

As described in 4.1.4, *ABCD1* is the major genetic determinant for X-ALD (Kemp *et al.*, 2011). Studies in X-ALD zebrafish models have revealed a crucial function of this candidate in myelinating glia, as *abcd1* mutants showed CNS hypomyelination and decreased numbers of oligodendrocytes due to increased apoptotic levels (Strachan *et al.*, 2017). Together, alterations detected for the *ABCD1 KO* in functional *in vitro* assays might partially correspond to this zebrafish phenotype. To mechanistically dissect the *ABCD1* relevance for this phenotype, it needs to be determined if VLCFAs accumulate at all in the *KO* clone. Investigations in X-ALD patient-derived iPSCs lacking ABCD1 have shown that oligodendrocytes did not display any phenotypic developmental delay suggesting ABC SUBFAMILY

MEMBER 2 (ABCD2) as compensatory transporter (Jang *et al.*, 2011). Members of the ABC family are half-size transporters, containing one transmembrane and one nucleotide binding domain and therefore require dimerization for proper functioning. Although they have been annotated to act predominantly as homodimers, an ABCD1-ABCD2 heterodimer has already been described and both family members have an overlapping substrate specificity (Geillon *et al.*, 2014; Morita and Imanaka, 2012). These findings are in line with *in vitro* results showing that only a CRISPR/Cas9 double KO of Abcd1 and Abcd2 resulted in increased VLCFAs levels in a microglial cell line (Raas *et al.*, 2019).
Taken all these facts into account, the milder phenotypic alterations of the differentiating *ABCD1 KO* clone might be explained as well by a compensatory effect of ABCD2 that is launched during neuronal cell fate induction. To validate this hypothesis, further investigations would be required.
Interestingly, in contrast to all other gene-specific *KO* clones, explorative qRT PCR data analysis displayed for many neuronal markers significantly increased relative mRNA expression levels in the *ABCD1 KO* clone compared to *mock control* clone. Moreover, the maximal relative expression level in the case of *MAP2, SYP* and *GAP43* was already reached after 7+14d of differentiation. Consequently, *ABCD1 KO* cells might either possess a premature neuronal-like maturation capacity or these enhanced expression levels might be caused by an overcompensated rescue mechanism of ABCD2 acting as complementary transporter. So far, IF analyses validated marker protein expression and confirmed respective subcellular localisations in differentiated *ABCD1 KO* cells. Nevertheless, for detection of mislocalised protein accumulations, higher imaging resolution technologies are necessary. Additionally, observations on mRNA level first need to be confirmed using a targeted analysis approach and not a purely descriptive statistical evaluation. Furthermore, investigations on protein level in a quantitative manner and expression profiling of ABCD2 should be included in the future.

In conclusion, functional *in vitro* analyses further strengthened a putative relevance of *ATP7A, SREBF1* and *ABCD1* for the development of HSCR, although none of the clones was completely disrupted in its neuronal maturation capacity.
qRT PCR results suggested that impairments in the candidate specific *KO* clones are not caused by a RET dependent mechanism. Remarkably, these analyses showed as well that all edited genes were still expressed in respective *KO* clones. This suggested that edited transcripts were not targeted by NMD. However, as outlined earlier, this cellular process was not investigated in the context of the presented project and mRNA half-life times must as well be taken into account for data interpretation (fig. 31).

Applying the presented experimental pipeline to a primary ENS cell population might help to dissect the functional relevance of *ATP7A, SREBF1, ABCD1* and *PIAS2* for ENS formation in detail and thereby as well uncover the underlying molecular pathways. For this purpose, enteric neuronal progenitor cells isolated from surgically resected gut specimens might be used (Metzger *et al.*, 2009; Schmitteckert *et al.*, 2019). Altered molecular signatures in the major ENS signalling pathways might be identified by expression profiling of respective *KO* clones using microarray or RNA Sequencing technologies.
Although the presented *in vitro* approach is proven feasible in the current study to analyse the putative functional relevance of a specific molecular player for HSCR, the model system does not reflect the

individual genetic background of the respective patient. This can only be achieved by using patient-derived cells. To finally proof the functional relevance of single HSCR candidate gene *in vitro*, gene-specific isogenic control lines would be required.

4.2.4 *In silico evidence for functional relevance of patient-specific gene variants*

All previously discussed findings add further evidence for the putative HSCR relevance of the selected candidates. However, the *in vitro* study pipeline of this project was based on a candidate specific loss-of-function approach. In contrast to this, studied HSCR patients carried only hetero- (*SREBF1, PIAS2)* or hemizygous *(ATP7A, ABCD1)* missense variants with unknown functional impact. To strengthen the suitability of the presented *in vitro* study pipeline, patient-specific variants were characterized on different levels (fig. 33).

In silico analyses demonstrated that - except for the *ABCD1* variant - none of the other mutations has so far been annotated in the European (non-Finnish) control cohort of the gnomAD dataset. The annotated MAF for the *ABCD1*-specific nucleotide exchange indicated a rare variant. Calculation of the CADD scores revealed that each variant can be classified as deleterious (cut off >13), as outlined previously.

On protein level, only the variant in PIAS2 mapped to a functionally annotated protein domain. Specifically, the missense variant (R>W) resided in the highly conserved PINIT region shown to be important for nuclear localisation and for substrate recognition (Duval *et al.*, 2003; Yunus and Lima, 2009). In contrast to this, ATP7A, SREBF1 and ABCD1-specific patient mutations did not reside in any functionally annotated domain.

The Arginine (charged) to Glutamine (uncharged) substitution in ATP7A was located C-terminally to the phosphorylation (P) domain, close to the second last transmembrane (TM) domain. This P-domain is crucial for ATP binding and autophosphorylation and thereby regulates the conformational change of the transporter molecule. In general, subcellular trafficking of ATP7A between the *trans* Golgi network and the plasma membrane is highly copper dependent and missense mutations have already been shown to affect correct protein localisation. Comparable to the HSCR patient-specific mutation, the majority of mutations causative for Menkes disease cluster at the C-terminal end of the protein (Kaler, 2011; Skjørringe *et al.*, 2017; Telianidis *et al.*, 2013).

The amino acid substitution in the case of SREBF1 resulted in an exchange of the negatively charged Aspartic Acid by the positively charged Histidine. As outlined previously, protein processing of SREBF1 is reflected by its distinct subcellular localisation either in the ER/the Golgi (precursor) or in the nucleus (mature) (Shao and Espenshade, 2014). Although proteolytic cleavage sites reside more to the N-terminus than the patient-specific amino acid substitution, the carboxy-terminus of SREBF1 has been shown to be crucial for proper protein maturation in a cholesterol-dependent manner (Horton *et al.*, 2002).

In the case of ABCD1, the patient-specific amino acid substitution (T (uncharged) > R (positively charged)) was located at the N-terminal end of the protein. A hydrophobic segment at this terminus has been shown to be crucial for correct subcellular localisation of the peroxisomal transporter and thus for proper VLCFA metabolism (Kawaguchi and Morita, 2016; Lee *et al.*, 2014a).

Protein localisations were investigated by IF analyses of cells, transiently transfected with recombinantly expressed WT and MUT candidate proteins. These analyses did not reveal any protein mislocalisation – neither in HEK293TN nor in SHSY5Y cells – caused by the specific mutations in any of the candidate proteins. Additionally, subcellular localisations of all WT and MUT proteins matched expectations as they confirmed previously described data.

However, several aspects need to be considered. First, recombinantly expressed candidate proteins contained specific tag sequences, which might interfere with proper subcellular trafficking and localisation in general. Second, analyses were solely carried out under standard cultivation conditions. However, in the case of ATP7A and SREBF1 it is well known that copper and sterol levels affect respective protein functions and localisations (Horton *et al.*, 2002; Skjørringe *et al.*, 2017). Third, no subcellular specific marker was co-stained with the recombinantly expressed candidate proteins. Conclusions are based only on observable staining patterns. And finally, variants might not impede correct subcellular protein localization, but might interfere with other candidate specific functions as e.g. VLCFA transport in the case of ABCD1 or substrate SUMOylation in the case of PIAS2. A putative impairment in the SUMOylation capacity of MUT PIAS2 might be assessed by immunoprecipitation and mass spectrometry analyses, followed by comparative data analysis. Additionally, the amino acid substitutions might be evaluated based on their impact for the respective protein tertiary structure, e.g. by *in silico* protein structure modeling. Altered tertiary structures might be causative for impaired protein functions.

In conclusion, while *in silico* analyses of all patient-specific candidate gene mutations suggested a functional relevance of the respective variants, no further *in vitro* evidence could be generated in the context of this project.

4.2.5 *Candidate gene confirmation based on risk assessment*

To evaluate and summarize all previously described and discussed findings, a scoring system was applied. To allow the most objective and unbiased evaluation of individual candidate genes, this strategy was based on a binary system.

Applying the scoring system to well established HSCR risk loci, validated its suitability for risk evaluation of novel HSCR candidate genes. Respective risk scores of *RET* (14), *SOX10* (13), *EDNRB* (12) and *GDNF* (12) were close to or even reached the maximum value (14) (Tab. 9). However, this system has also limitations which are uncovered by the assessment of *PHOX2B*. Although this gene is well characterized for its role in ENS development and for its relevance in HSCR, no rare disease-associated mutations could be identified in literature. Hence, no scores for *PHOX2B* could be attributed in the ‘Gene mutation‘ category for all prediction tools ($\Sigma_1=\Sigma_2$) (Bachelor Thesis Dawid, 2019).

As outlined in 2.2.1, scores within the ‘Gene expression‘ category were mainly based on evidences annotated in different databases. Consequently, missing database entries or incomplete data compilation may introduce a bias in the evaluation of specific genes – as illustrated for *PHOX2B* and *ABCD1*. No protein expression data was available for the candidate gene in the Human Protein Atlas, although ABCD1 expression in human colon and brain has already been investigated (Höftberger *et al.*, 2007; The Human Protein Atlas, ABCD1). Nevertheless, all four selected candidate genes displayed

rather high Σ_1 and Σ_2 scores, which confirmed the initial candidate gene selection procedure in a reproducible and robust manner. In the future, this scoring system will be used to facilitate the selection of novel and most promising HSCR candidate genes identified by exome sequencing of additional patients.
Evidence implemented in the calculation of Σ_2 were combined with functional *in vitro* data of this project. This further strengthened the putative role of *ATP7A, SREBF1* and *ABCD1* in the HSCR aetiology, as respective Σ_3 values were highly comparable to the risk score of the major HSCR disease loci *RET*.

In conclusion, the complementary research approach presented in this study could be evaluated in an objective manner through application of the presented scoring system. Modifications might be applied to this system in the future to further improve its validity. To this end, additional categories might be introduced or classifications and cut offs might be adjusted based on available data and current knowledge.

CONCLUSIONS AND OUTLOOK

This project aimed to establish a complementary research pipeline for the elucidation of novel candidate genes in HSCR, taking genetic, bioinformatics, molecular, cellular and functional data into account. Application of this strategy in an exemplary way to two sporadic L-HSCR patients led to the identification of four novel HSCR risk loci: *ATP7A*, *SREBF1*, *ABCD1* and *PIAS2*. This classification was confirmed based on the calculation of a single risk score, which summarized various evidences generated in this project and thereby underscored the suitability and strength of a complementary research approach in the context of a rare and complex disease as HSCR.

The complex genetics of HSCR in correlation with respective genetic findings of this study suggest that *ATP7A, SREBF1, ABCD1* and *PIAS2* probably contribute to the disease. A variable number of further genetic risk variants might be present in both affected individuals.

Interestingly, on molecular level all selected candidates share one similarity: protein functions of ATP7A, SREBF1, ABCD1 and PIAS2 are not highly specific to a particular cell type but are rather of general cellular importance. This was as well reflected by the respective expression patterns in various tissues. Most likely, these novel risk factors might contribute to the development of the HSCR phenotype at various levels rather than affecting only a specific cell type, cell function or signalling pathway. This hypothesis was further supported by functional *in vitro* characterization of gene-specific *KO* cell clones and by *in silico* evidences based on the IPA knowledge base. Furthermore, so far unknown HSCR risk loci in both patients might contribute in the same way thereby shaping the individual's phenotype. This conclusion is in line with knowledge from recent studies, where complex interactions between multiple risk loci were suggested to be crucial for the individual's HSCR risk.

Generation of more detailed knowledge about the molecular mechanisms of HSCR is pivotal in order to understand its postoperative complications and to change the current treatment regime in future. The presented study underscored the necessity and strength of a complementary research approach, but as well highlighted major challenges and future requirements, which need to be overcome in the field of HSCR research.

SUPPLEMENTARY

Genetic evaluation

WES data re-evaluation

Table S1: Filtered WES data of family I
CADD scores were calculated using the CADD model GRCh37-v1.4 (https://cadd.gs.washington.edu/snv). *For indels, local CADD_PHRED values were calculated by Dr. Nagarajan Paramasivam (Theoretical Bioinformatics Division, DKFZ).

Gene name	Nucleotide/Amino acid change	Mutation type	CADD
ATP7A	ENST00000350425.4 exon8:c.G1076A:p.R359Q	hemizygous	26.9
DNAH14	ENST00000439375.2 exon29:c.4857_4860del:p.1619_1620del	compound heterozygous	36*
DNAH14	ENST00000445597.2 exon3:c.G356A:p.G119D,	compound heterozygous	8.636
DSG1	ENST00000462981.2 exon4:c.C192T:p.Y64Y,	compound heterozygous	23.6
DSG1	ENST00000462981.2 exon4:c.G685A:p.G229S	compound heterozygous	24
KIAA1522	ENST00000401073.2 exon1:c.G68T:p.R23L,	*de novo*	23.2
PHF16	ENST00000424392.1 exon5:c.C409T:p.R137C	hemizygous	26.9
PTPRK	ENST00000368227.3 exon17:c.G2564A:p.R855H	compound heterozygous	22.6
PTPRK	ENST00000368227.3 exon12:c.G1929T:p.K643N	compound heterozygous	25.8
PVRL2	ENST00000252485.4 exon2:c.C227T:p.P76L	*de novo*	12.23
SREBF1	ENST00000355815.4 exon19:c.G3262C:p.D1088H,	compound heterozygous	25.5
SREBF1	ENST00000355815.4 exon14:c.G2525A:p.R842Q,	compound heterozygous	18.62
STAG2	ENST00000455404.1 exon3:c.A116G:p.K39R	hemizygous	23.5
STRC	ENST00000541030.1 exon25:c.C2599T:p.L867F	compound heterozygous	24.7
STRC	ENST00000541030.1 exon23:c.C2242T:p.R748W	compound heterozygous	23.7
ZNF275	ENST00000421401.3 exon4:c.T877C:p.Y293H	hemizygous	23.5

Table S2: Filtered WES data of family II
CADD scores were calculated using the CADD model GRCh37-v1.4 (https://cadd.gs.washington.edu/snv).

Gene name	Nucleotide/Amino acid change	Mutation type	CADD
ABCD1	ENST00000218104.3 exon1:c.C41G:p.T14R	hemizygous	15.38
CNKSR2	ENST00000425654.2 exon19:c.A2588C:p.N863T	hemizygous	18.93
EBP	ENST00000495186.1 exon5:c.C511T:p.R171C,	hemizygous	21.6
FER1L6	ENST00000522917.1 exon22:c.G2827T:p.G943W	compound heterozygous	33
FER1L6	ENST00000522917.1 exon23:c.T2981G:p.V994G	compound heterozygous	28.4
MXRA5	ENST00000217939.6 exon5:c.G1783A:p.V595M	hemizygous	15.17
OR5P2	ENST00000329434.2 exon1:c.G56A:p.G19D,	hemizygous	23.7
PIAS2	ENST00000585916.1 exon7:c.C889T:p.R297W	*de novo*	26.9
RNF213	ENST00000456466.1 exon9:c.C1637T:p.T546I	compound heterozygous	8.843
RNF213	ENST00000582970.1 exon29:c.A6979G:p.N2327D	compound heterozygous	4.350
SLC2A4	ENST00000571308.1 exon3:c.C236G:p.T79S	homozygous	14.96
SPTA1	ENST00000368147.4 exon50:c.G6927T:p.M2309I	compound heterozygous	25.2
SPTA1	ENST00000368147.4 exon32:c.A4564G:p.T1522A	compound heterozygous	17.61
SYNE2	ENST00000394768.2 exon33:c.T4907C:p.I1636T	compound heterozygous	16.08
SYNE2	ENST00000441438.2 exon4:c.A488T:p.Q163L	compound heterozygous	14.60
TENM4	ENST00000278550.7 exon31:c.C5729T:p.A1910V	*de novo*	21.3
TFDP3	ENST00000310125.4 exon1:c.G427A:p.A143T	hemizygous	7.926
VSIG4	ENST00000412866.2 exon7:c.C866T:p.T289I	hemizygous	13.84
VSIG4	ENST00000455586.2 exon2:c.G274T:p.V92F	hemizygous	14.96

Summary and overview

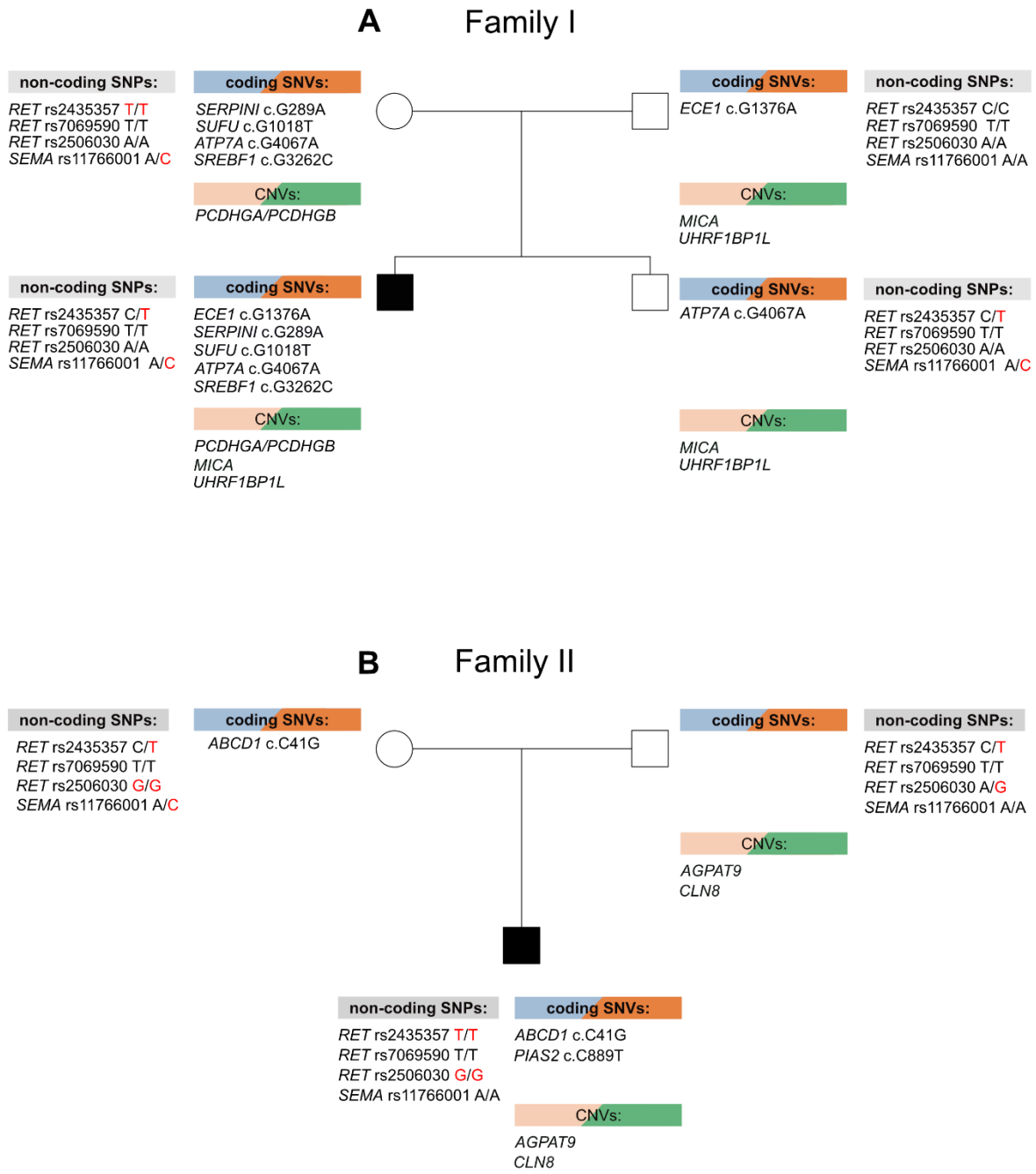


Figure S1: Genotype overview on selected non-coding and coding genetic variants for all individuals of family I and II

A and B: More detailed information about the coding SNVs and CNVs are shown in section 3.1.

IPA analysis

Table S3: ENS and HSCR relevant genes used for IPA analysis and for comparison of rare SNVs
(based on (Gui *et al.*, 2017; Lake and Heuckeroth, 2013; Luzón-Toro *et al.*, 2015b))

Gene	Gene ID	Gene	Gene ID	Gene	Gene ID
ALDH1A2	ENSG00000128918	*HOXA4*	ENSG00000197576	*RBP4*	ENSG00000138207
ARHGEF3	ENSG00000163947	*HOXB5*	ENSG00000120075	*RET*	ENSG00000165731
ARTN	ENSG00000117407	*HOXD4*	ENSG00000170166	*RMRP*	ENSG00000277027
ASCL-1	ENSG00000139352	*IFNGR2*	ENSG00000159128	*SALL4*	ENSG00000101115
BBS1-11	ENSG00000174483	*IHH*	ENSG00000163501	*SCG3*	ENSG00000104112
CADM1	ENSG00000182985	*IKBKAP*	ENSG00000070061	*SEMA3A*	ENSG00000075213
CARTPT	ENSG00000164326	*IL10RB*	ENSG00000243646	*SEMA3C*	ENSG00000075223
CBR1	ENSG00000159228	*ITGB1*	ENSG00000150093	*SEMA3D*	ENSG00000153993
CDC42	ENSG00000070831	*JAG1*	ENSG00000101384	*SERPINI1*	ENSG00000163536
CDH2	ENSG00000170558	*JAG2*	ENSG00000184916	*SHH*	ENSG00000164690
CRMP1	ENSG00000072832	*KIAA1279*	ENSG00000198954	*SLC6A2*	ENSG00000103546
CSTB	ENSG00000160213	*KIF26A*	ENSG00000066735	*SMO*	ENSG00000128602
CTNNAL1	ENSG00000119326	*KLF4*	ENSG00000136826	*SOD1*	ENSG00000142168
DCC	ENSG00000187323	*L1CAM*	ENSG00000198910	*SON*	ENSG00000159140
DCX	ENSG00000077279	*LGI4*	ENSG00000153902	*SOX10*	ENSG00000100146
DHCR7	ENSG00000172893	*MAB21L1*	ENSG00000180660	*SOX2*	ENSG00000181449
DICER1	ENSG00000100697	*MAPK10*	ENSG00000109339	*SOX8*	ENSG00000005513
DLL1	ENSG00000198719	*MAPT*	ENSG00000186868	*SPRY2*	ENSG00000136158
DLL3	ENSG00000090932	*MLLT11*	ENSG00000213190	*STMN2*	ENSG00000104435
DLX1	ENSG00000144355	*NAV2*	ENSG00000166833	*STMN3*	ENSG00000197457
DNMT3B	ENSG00000088305	*NKX2-1*	ENSG00000136352	*SUFU*	ENSG00000107882
DPYSL3	ENSG00000113657	*NOG*	ENSG00000183691	*SYT11*	ENSG00000132718
EBF3	ENSG00000108001	*NOTCH1*	ENSG00000148400	*TAGLN3*	ENSG00000144834
ECE1	ENSG00000117298	*NOTCH2*	ENSG00000134250	*TBX3*	ENSG00000135111
EDN3	ENSG00000124205	*NOTCH3*	ENSG00000074181	*TCF4*	ENSG00000196628
EDNRB	ENSG00000136160	*NRG1*	ENSG00000157168	*TCF7L2*	ENSG00000148737
ELAVL2	ENSG00000107105	*NRG3*	ENSG00000185737	*TCOF1*	ENSG00000070814
ELAVL4	ENSG00000162374	*NRP1*	ENSG00000099250	*TFAM*	ENSG00000108064
ERBB2	ENSG00000141736	*NRTN*	ENSG00000171119	*TFF3*	ENSG00000160180
ERBB3	ENSG00000065361	*NTF3*	ENSG00000185652	*TGFB2*	ENSG00000092969
ERBB4	ENSG00000178568	*NTRK3*	ENSG00000140538	*TLX2*	ENSG00000115297
ERCC1	ENSG00000012061	*PAX3*	ENSG00000135903	*TMEFF2*	ENSG00000144339
ETV1	ENSG00000006468	*PCDHA1*	ENSG00000204970	*TPH2*	ENSG00000139287
FGF13	ENSG00000129682	*PDS5A*	ENSG00000121892	*TREX1*	ENSG00000213689
GAP43	ENSG00000172020	*PDS5B*	ENSG00000083642	*TTC3*	ENSG00000182670
GDNF	ENSG00000168621	*PFKL*	ENSG00000141959	*TUBB3*	ENSG00000258947
GFRA1	ENSG00000151892	*PHACTR4*	ENSG00000204138	*UCHL1*	ENSG00000154277
GFRA2	ENSG00000168546	*PHOX2A*	ENSG00000165462	*VIP*	ENSG00000146469
GFRA3	ENSG00000146013	*PHOX2b*	ENSG00000109132	*ZEB2*	ENSG00000169554
GFRA4	ENSG00000125861	*PLXNA1*	ENSG00000114554	*ZIC2*	ENSG00000043355
GLI1	ENSG00000111087	*PLXNB1*	ENSG00000164050		
GLI2	ENSG00000074047	*POFUT1*	ENSG00000101346		
GLI3	ENSG00000106571	*PROK1*	ENSG00000143125		
GNG2	ENSG00000186469	*PROK2*	ENSG00000163421		
GNG3	ENSG00000162188	*PROKR1*	ENSG00000169618		
GRB10	ENSG00000106070	*PROKR2*	ENSG00000101292		
HAND2	ENSG00000164107	*PRPH*	ENSG00000135406		
HES1	ENSG00000114315	*PSPN*	ENSG00000125650		
HLX	ENSG00000136630	*PTCH1*	ENSG00000185920		
HMP19	ENSG00000170091	*PTEN*	ENSG00000171862		
HMX3	ENSG00000188620	*RAC1*	ENSG00000136238		

Database research of candidate genes

Table S4: Overview on annotated protein coding isoforms for candidates in different databases
Transcript identifiers are given.

	ATP7A	***SREBF1***	***ABCD1***	***PIAS2***
Ensembl	ENST00000341514.11 ENST00000343533.9 ENST00000645454.1 ENST00000642651.1	ENST00000355815.8 ENST00000261646.9 ENST00000395757.5 ENST00000423161.3 ENST00000478616.1 ENST00000577897.1 ENST00000486311.5	ENST00000218104.6 ENST00000370129.4	ENST00000585916.6 ENST00000324794.11 ENST00000592212.5 ENST00000545673.5 ENST00000589917.1 ENST00000592221.5 ENST00000586953.1 ENST00000587810.1
NCBI	NM_000052.7 NM_001282224.1	NM_001321096.2 NM_004176.4 NM_001005291.2	NM_000033.4	NM_004671.5 NM_001354034.1 NM_001324058.1 NM_173206.3 NM_001324046.1 NM_001324048.1 NM_001324059.1 NM_001324051.1 NM_001324055.1 NM_001324052.1 NM_001324054.1 NM_001324047.1 NM_001324057.1 NM_001324053.1 NM_001324049.1 NM_001354038.1 NM_001354039.1 NM_001354036.1 NM_001354033.1 NM_001354035.1 NM_001354037.1 NM_001324060.1

Negative control staining

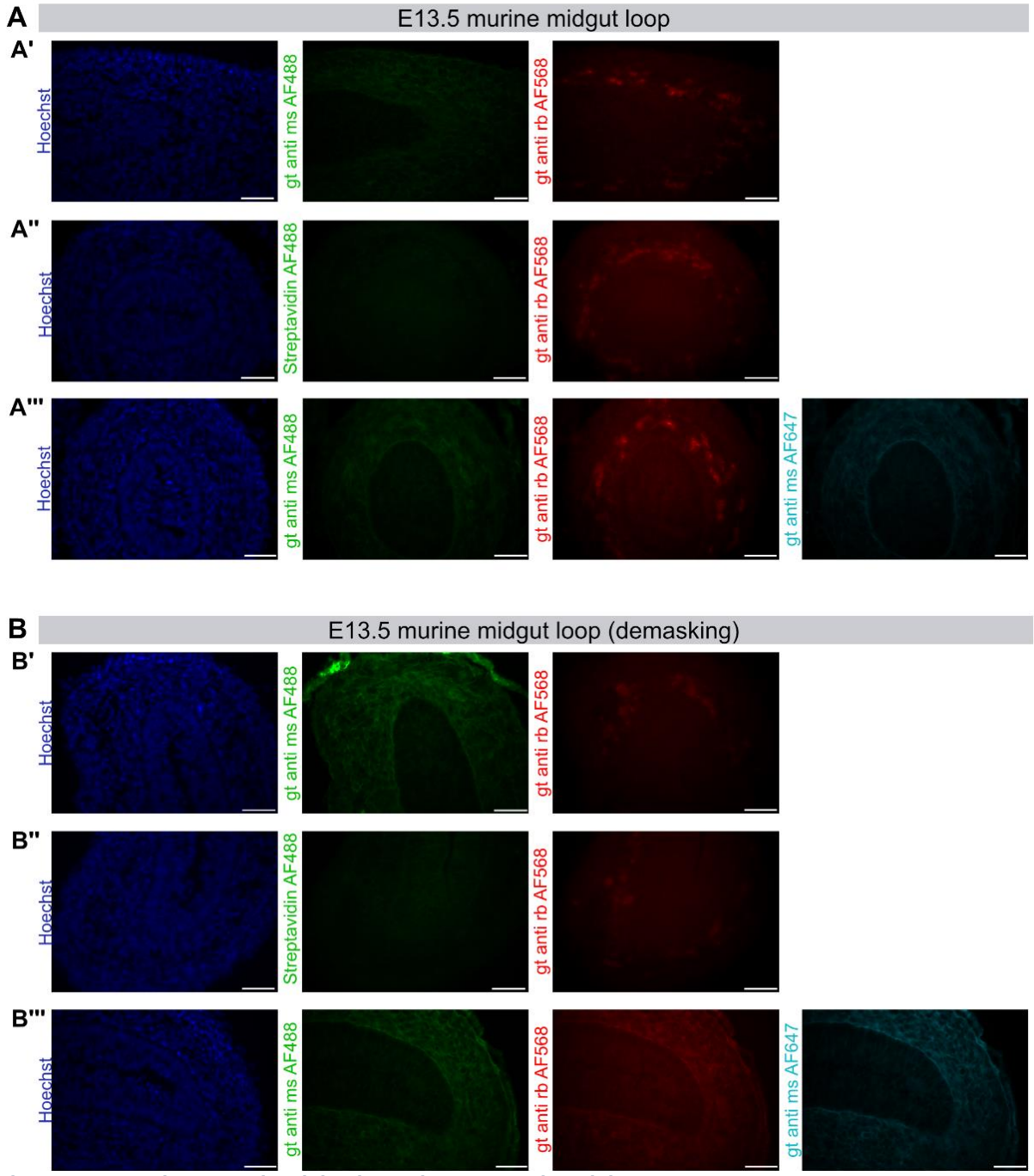


Figure S2: Negative control staining in murine cryosectioned tissue at E13.5

A and B: Control staining for different secondary antibody combinations (A'/B': goat anti mouse AF488 (green), goat anti mouse AF568 (red); A''/B'': Streptavidin AF488 (green), goat anti rabbit AF568 (red); A'''/B''': goat anti mouse AF488 (green), goat anti rabbit AF568 (red), goat anti mouse AF647 (turquoise). B: Tissue was demasked before IF staining. Nuclei were counterstained with Hoechst (blue). Scale bar: 50 µm.

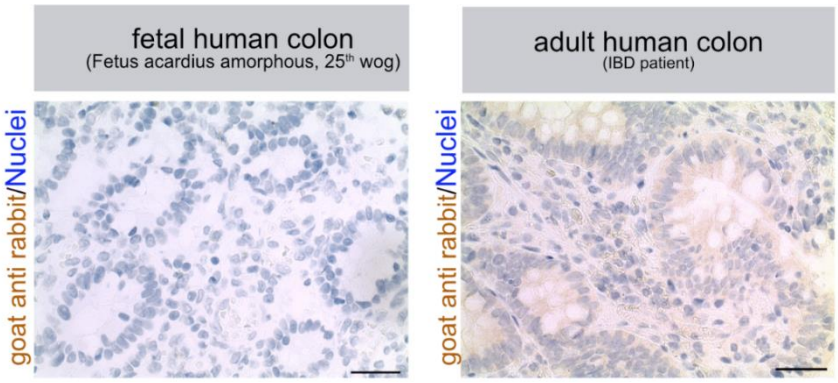


Figure S3: IHC negative control staining in human FFPE sections
Control staining for peroxidase conjugated AffiniPure goat anti rabbit antibody (brown) in fetal and adult human colon sections were performed by Dr. Cornelia Thöni and Jutta Scheuerer (research group of Dr. Felix Lasitschka, Institute of Pathology, University Hospital Heidelberg). Images are representatives. Nuclei were counterstained with Hematoxylin (blue). Scale bar: 20 µm.

CRISPR/Cas9 genome editing

Endogenous candidate gene expression

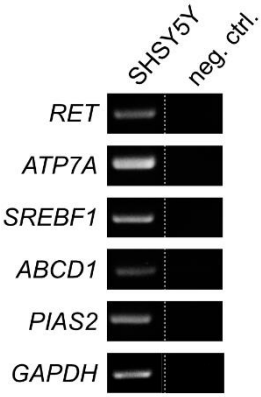


Figure S4: Endogenous gene expression in WT SHSY5Y cells
Results for conventional PCR analyses are shown. *GAPDH* served as reference.

sgRNA evaluation

Table S5: sgRNA evaluation using different design portals
Ranks are listed. *no longer available

Gene	Exon2	sgRNA sequence (5>3)	Ranking		
			CCTop	*CHOPCHOP*	*Optimized CRISPR design portal**
RET	Ex2	TCTACGGCACGTACCGCACA TGTGCGGTACGTGCCGTAGA	1	-	1
ATP7A	Ex3	GATCATAGTATGGCTCAAGC GCTTGAGCCATACTATGATC	2	15	-
SREBF1	Ex5	GAGACCTGCCGCCTTCACAG CTGTGAAGGCGGCAGGTCTC	1	2	3
ABCD1	Ex2	CGCTCCACACATACTTCATG CATGAAGTATGTGTGGAGCG	2	1	11
PIAS2	Ex6	TAATATTCAAGGGGCGTCCA TGGACGCCCCTTGAATATTA	1	2	1
	Ex2	CGATATCCACGAACTCTTGA TCAAGAGTTCGTGGATATCG	2	8	3
		CCGAGAATTGTATAGACGCCGAT ATCGGCGTCTATACAATTCTCGG	3	-	2
		CCACGAACTCTTGAAGGACTTTC GAAAGTCCTTCAAGAGTTCGTGG	4	-	7

sgRNA cloning

Table S6: Thermal cycler program used for phosphorylation and annealing of sgRNA oligonucleotides

step	temperature	time
1	37°C	30 min
2	95°C	5 min
ramp down to 25°C at 5°C/min		

KO validation by Sanger Sequencing

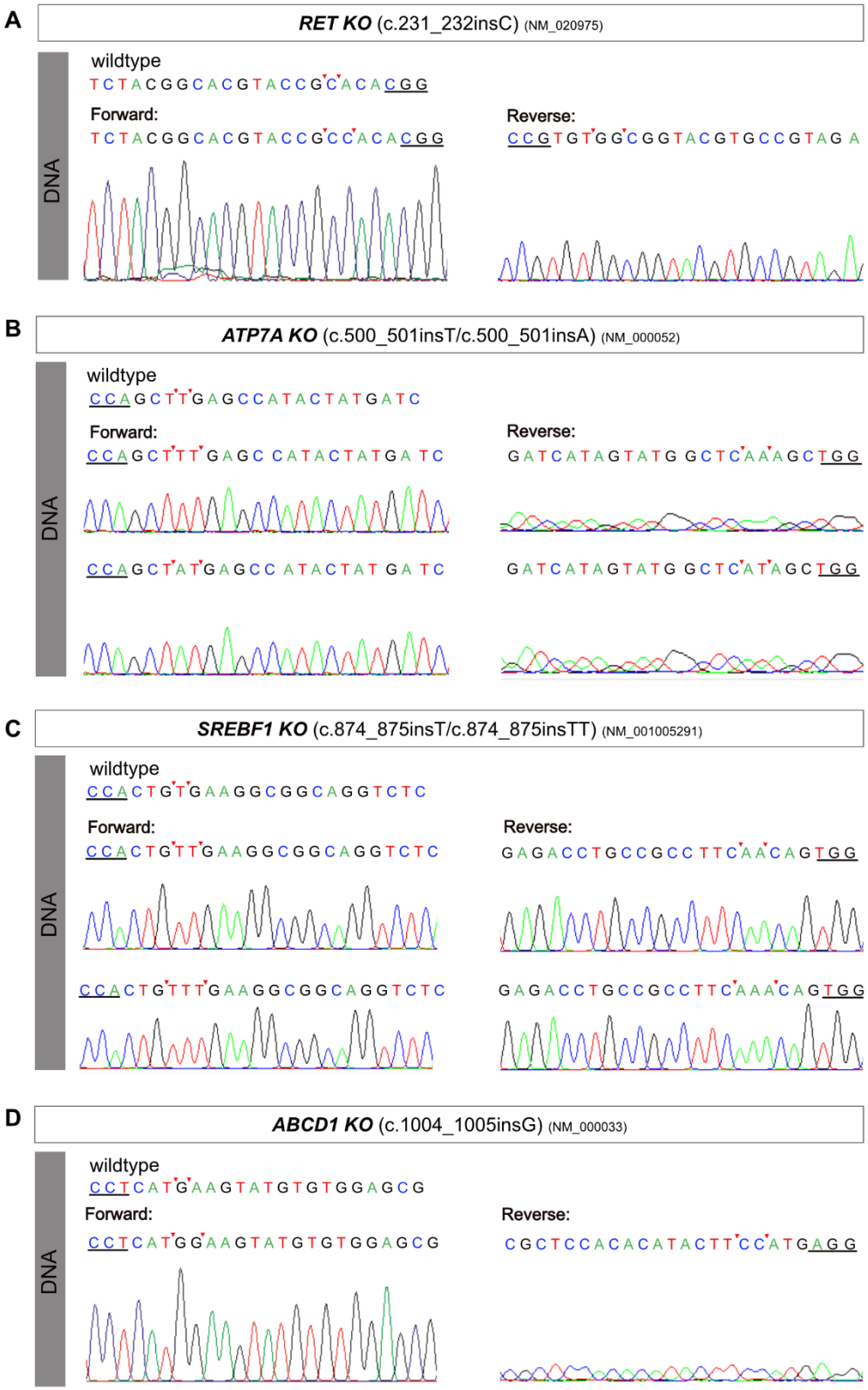


Figure S5: Forward and reverse Sanger sequencing of CRISPR/Cas9 engineered KO clones

A-D: CRISPR/Cas9 mediated genome modifications (homozygous or compound heterozygous) were verified by forward and reverse Sanger sequencing. Gene-specific sgRNA target sites and four coloured chromatograms are shown. NCBI transcript identifiers used for annotation of the editing event are displayed. Probable Cas9 cutting sites are highlighted by red triangles. PAM sites are underscored.

Analysis of alternative splicing in KO clones

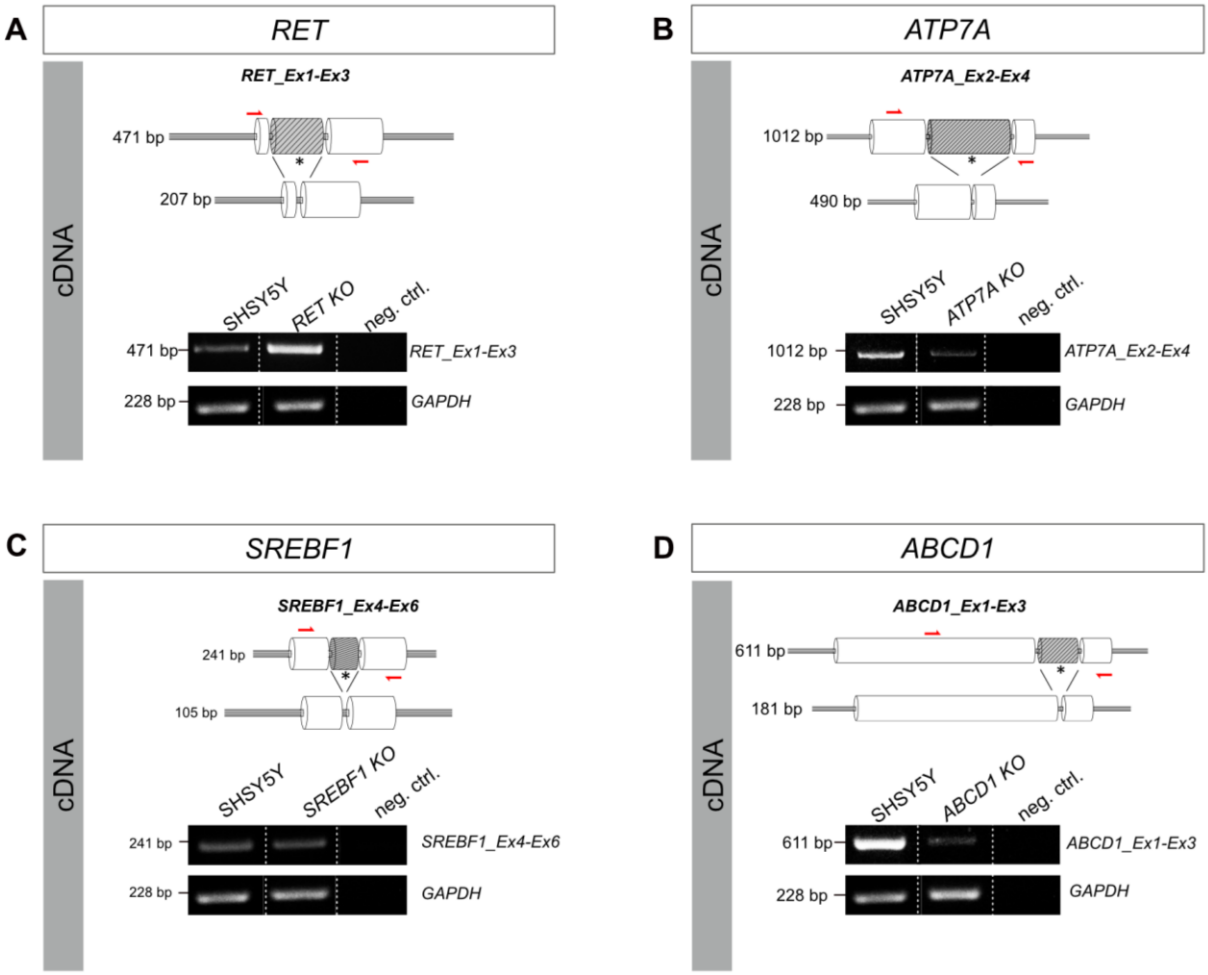


Figure S6: Analysis of alternative splicing of the genome modified exon in individual KO clones
A-D: In all genome edited *KO* clones, alternative splicing of the genetically modified exon (highlighted by an asterisk) was excluded by conventional PCR analyses on cDNA level. Primers are illustrated by red arrows. GAPDH was used as reference. Expected and observed amplicon sizes are annotated.

Off-target analysis

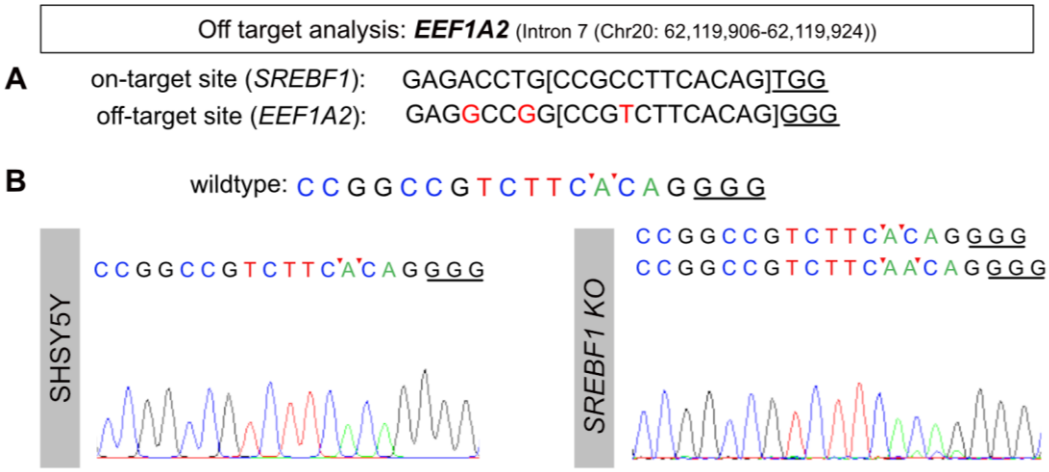


Figure S7: Off-target effect in the SREBF1 KO clone

A: On-target and off-target sequence of the *SREBF1* specific sgRNA. PAM site is underscored. Mismatching nucleotides are highlighted in red. The core sequence close to the PAM (underscored) is marked. **B:** An unintended heterozygous one nucleotide insertion was detected in the intronic region of *EEF1A2* in the *SREBF1 KO*. This indel was located at the respective position of the PAM and was not present in the parental SHSY5Y line. PAM sites are underscored. Putative Cas9 cutting sites are highlighted by red triangles. Four colored chromatograms from Sanger sequencing are shown.

qRT PCR

Table S7: Thermal cycler program for qRT PCR

step	temperature	time	
1	95° C	2 min	
2	95° C 60° C	5 sec 20 sec	X 40
3	Melting curve		

Vector maps

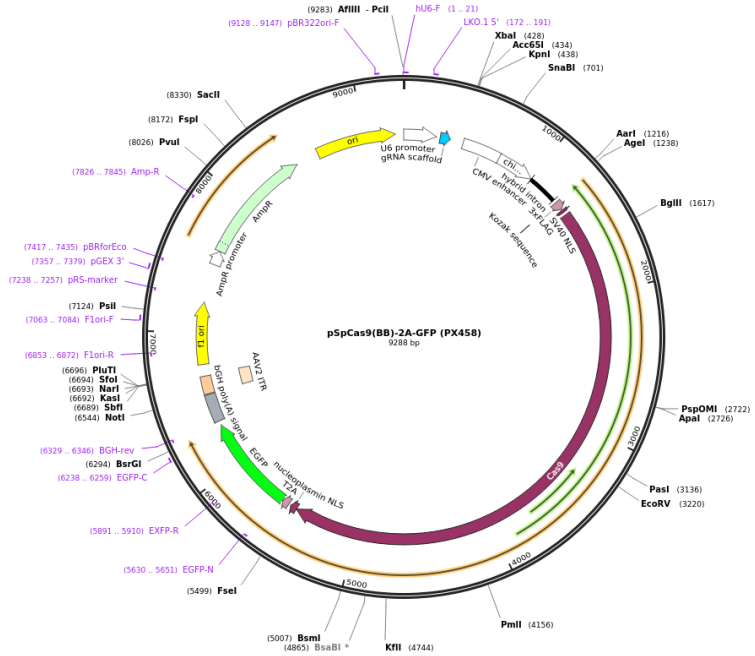


Figure S8:* pSpCas*(BB)-2A-GFP vector map
(from: https://www.addgene.org/48138/)

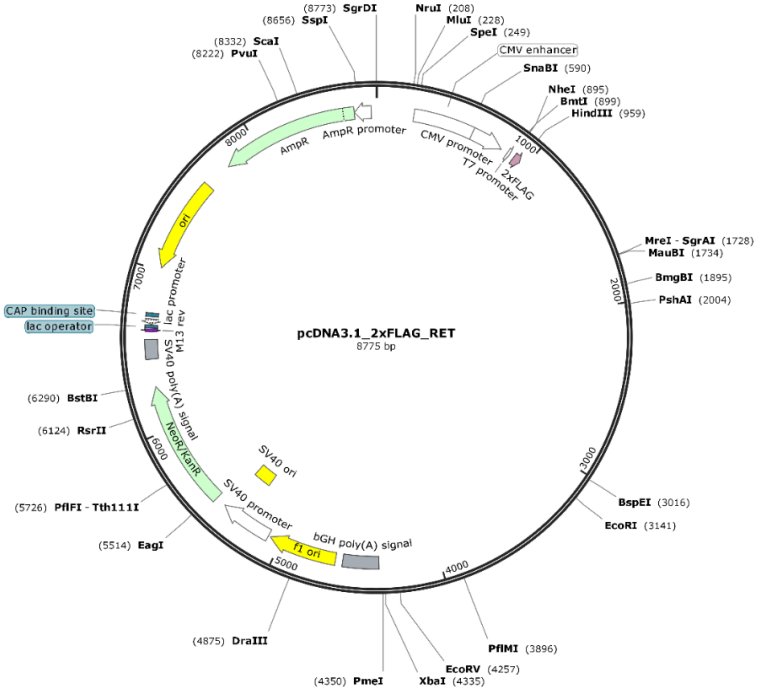


Figure S9: pcDNA3.1_2xFLAG_RET vector map

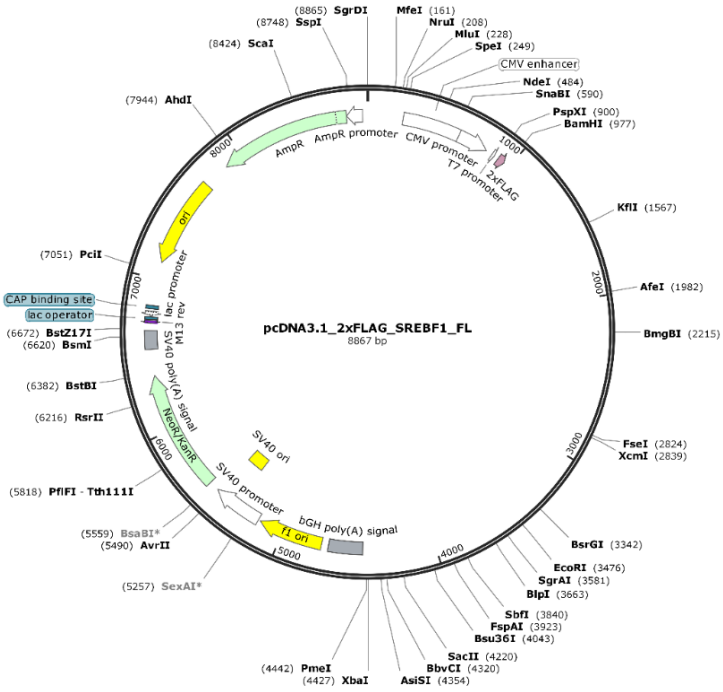


Figure S10: pcDNA3.1_2XFLAG_SREBF1B_FL vector map

LITERATURE

Acosta, S., Lavarino, C., Paris, R., Garcia, I., de Torres, C., Rodríguez, E., Beleta, H., Mora, J., 2009, "*Comprehensive characterization of neuroblastoma cell line subtypes reveals bilineage potential similar to neural crest stem cells*", BMC developmental biology 9, 12.

Addgene, 2017, "*CRISPR 101: A Desktop Resource*", 2nd ed. www.addgene.org.

Agholme, L., Lindstrom, T., Kagedal, K., Marcusson, J., Hallbeck, M., 2010, "*An in vitro model for neuroscience: differentiation of SH-SY5Y cells into cells with morphological and biochemical characteristics of mature neurons*", J Alzheimers Dis 20, 1069-1082.

Amiel, J., Sproat-Emison, E., Garcia-Barcelo, M., Lantieri, F., Burzynski, G., Borrego, S., Pelet, A., Arnold, S., Miao, X., Griseri, P., Brooks, A.S., Antinolo, G., de Pontual, L., Clement-Ziza, M., Munnich, A., Kashuk, C., West, K., Wong, K.K., Lyonnet, S., Chakravarti, A., Tam, P.K., Ceccherini, I., Hofstra, R.M., Fernandez, R., Hirschsprung Disease, C., 2008, "*Hirschsprung disease, associated syndromes and genetics: a review*", J Med Genet 45, 1-14.

Anderson, D.B., Wilkinson, K.A., Henley, J.M., 2009, "*Protein sumoylation in neuropathological conditions*", Drug News Perspect 22, 255-265.

Andreoletti, P., Raas, Q., Gondcaille, C., Cherkaoui-Malki, M., Trompier, D., Savary, S., 2017, "*Predictive Structure and Topology of Peroxisomal ATP-Binding Cassette (ABC) Transporters*", International journal of molecular sciences 18.

Angrist, M., Bolk, S., Halushka, M., Lapchak, P.A., Chakravarti, A., 1996, "*Germline mutations in glial cell line-derived neurotrophic factor (GDNF) and RET in a Hirschsprung disease patient*", Nature genetics 14, 341-344.

Ashino, T., Sudhahar, V., Urao, N., Oshikawa, J., Chen, G.F., Wang, H., Huo, Y., Finney, L., Vogt, S., McKinney, R.D., Maryon, E.B., Kaplan, J.H., Ushio-Fukai, M., Fukai, T., 2010, "*Unexpected role of the copper transporter ATP7A in PDGF-induced vascular smooth muscle cell migration*", Circulation research 107, 787-799.

Autism Spectrum Disorders Working Group of the Psychiatric Genomics Consortium, 2017, "*Meta-analysis of GWAS of over 16,000 individuals with autism spectrum disorder highlights a novel locus at 10q24.32 and a significant overlap with schizophrenia*", Molecular autism 8, 21.

Bahrami, A., Joodi, M., Moetamani-Ahmadi, M., Maftouh, M., Hassanian, S.M., Ferns, G.A., Avan, A., 2017, "*Genetic Background of Hirschsprung Disease: A Bridge Between Basic Science and Clinical Application*", J Cell Biochem.

Baker, A., Carrier, D.J., Schaedler, T., Waterham, H.R., van Roermund, C.W., Theodoulou, F.L., 2015, "*Peroxisomal ABC transporters: functions and mechanism*", Biochemical Society transactions 43, 959-965.

Barlow, A., de Graaff, E., Pachnis, V., 2003, "*Enteric nervous system progenitors are coordinately controlled by the G protein-coupled receptor EDNRB and the receptor tyrosine kinase RET*", Neuron 40, 905-916.

Belkind-Gerson, J., Graeme-Cook, F., Winter, H., 2006, "*Enteric nervous system disease and recovery, plasticity, and regeneration*", J Pediatr Gastroenterol Nutr 42, 343-350.

Bell, N., Hann, V., Redfern, C.P., Cheek, T.R., 2013, "*Store-operated Ca(2+) entry in proliferating and retinoic acid-differentiated N- and S-type neuroblastoma cells*", Biochimica et biophysica acta 1833, 643-651.

Benarroch, E.E., 2007, "*Enteric nervous system: functional organization and neurologic implications*", Neurology 69, 1953-1957.

Bhatt, S., Diaz, R., Trainor, P.A., 2013, "*Signals and switches in Mammalian neural crest cell differentiation*", Cold Spring Harbor perspectives in biology 5.

Biedler, J.L., Helson, L., Spengler, B.A., 1973, "*Morphology and growth, tumorigenicity, and cytogenetics of human neuroblastoma cells in continuous culture*", Cancer research 33, 2643-2652.

Bondurand, N., Dufour, S., Pingault, V., 2018, "*News from the endothelin-3/EDNRB signaling pathway: Role during enteric nervous system development and involvement in neural crest-associated disorders*", Developmental biology.

Bondurand, N., Southard-Smith, E.M., 2016, "*Mouse models of Hirschsprung Disease and other developmental disorders of the Enteric Nervous System: Old and new players*", Developmental biology 417, 139-157.

Bunone, G., Borrello, M.G., Picetti, R., Bongarzone, I., Peverali, F.A., de Franciscis, V., Della Valle, G., Pierotti, M.A., 1995, "*Induction of RET proto-oncogene expression in neuroblastoma cells precedes neuronal differentiation and is not mediated by protein synthesis*", Experimental cell research 217, 92-99.

Burke, R., Commons, E., Camakaris, J., 2008, "*Expression and localisation of the essential copper transporter DmATP7 in Drosophila neuronal and intestinal tissues*", Int J Biochem Cell Biol 40, 1850-1860.

Burn, B., Brown, S., Chang, C., 2011, "*Regulation of early Xenopus development by the PIAS genes*", Dev Dyn 240, 2120-2126.

Burns, A.J., Goldstein, A.M., Newgreen, D.F., Stamp, L., Schafer, K.H., Metzger, M., Hotta, R., Young, H.M., Andrews, P.W., Thapar, N., Belkind-Gerson, J., Bondurand, N., Bornstein, J.C., Chan, W.Y., Cheah, K., Gershon, M.D., Heuckeroth, R.O., Hofstra, R.M., Just, L., Kapur, R.P., King, S.K., McCann, C.J., Nagy, N., Ngan, E., Obermayr, F., Pachnis, V., Pasricha, P.J., Sham, M.H., Tam, P., Vanden Berghe, P., 2016, "*White paper on guidelines concerning enteric nervous system stem cell therapy for enteric neuropathies*", Developmental biology 417, 229-251.

Calkins, C.M., 2018, "*Hirschsprung Disease beyond Infancy*", Clinics in colon and rectal surgery 31, 51-60.

Camargo, N., Smit, A.B., Verheijen, M.H., 2009, "*SREBPs: SREBP function in glia-neuron interactions*", The FEBS journal 276, 628-636.

Cantrell, V.A., Owens, S.E., Chandler, R.L., Airey, D.C., Bradley, K.M., Smith, J.R., Southard-Smith, E.M., 2004, "*Interactions between Sox10 and EdnrB modulate penetrance and severity of aganglionosis in the Sox10Dom mouse model of Hirschsprung disease*", Human molecular genetics 13, 2289-2301.

Cermenati, G., Audano, M., Giatti, S., Carozzi, V., Porretta-Serapiglia, C., Pettinato, E., Ferri, C., D'Antonio, M., De Fabiani, E., Crestani, M., Scurati, S., Saez, E., Azcoitia, I., Cavaletti, G., Garcia-Segura, L.M., Melcangi, R.C., Caruso, D., Mitro, N., 2015, "*Lack of sterol regulatory element binding factor-1c imposes glial Fatty Acid utilization leading to peripheral neuropathy*", Cell metabolism 21, 571-583.

Chalazonitis, A., Rothman, T.P., Chen, J., Gershon, M.D., 1998, "*Age-dependent differences in the effects of GDNF and NT-3 on the development of neurons and glia from neural crest-derived precursors immunoselected from the fetal rat gut: expression of GFRalpha-1 in vitro and in vivo*", Developmental biology 204, 385-406.

Chatterjee, S., Kapoor, A., Akiyama, J.A., Auer, D.R., Lee, D., Gabriel, S., Berrios, C., Pennacchio, L.A., Chakravarti, A., 2016, "*Enhancer Variants Synergistically Drive Dysfunction of a Gene Regulatory Network In Hirschsprung Disease*", Cell 167, 355-368.e310.

Chen, D., Tang, J.X., Li, B., Hou, L., Wang, X., Kang, L., 2018, "*CRISPR/Cas9-mediated genome editing induces exon skipping by complete or stochastic altering splicing in the migratory locust*", BMC biotechnology 18, 60.

Chen, Y., Bang, S., McMullen, M.F., Kazi, H., Talbot, K., Ho, M.X., Carlson, G., Arnold, S.E., Ong, W.Y., Kim, S.F., 2017a, "*Neuronal Activity-Induced Sterol Regulatory Element Binding Protein-1 (SREBP1) is Disrupted in Dysbindin-Null Mice-Potential Link to Cognitive Impairment in Schizophrenia*", Molecular neurobiology 54, 1699-1709.

Chen, Z., Donnelly, C.R., Dominguez, B., Harada, Y., Lin, W., Halim, A.S., Bengoechea, T.G., Pierchala, B.A., Lee, K.F., 2017b, "*p75 Is Required for the Establishment of Postnatal Sensory Neuron Diversity by Potentiating Ret Signaling*", Cell reports 21, 707-720.

Ciccarone, V., Spengler, B.A., Meyers, M.B., Biedler, J.L., Ross, R.A., 1989, "*Phenotypic diversification in human neuroblastoma cells: expression of distinct neural crest lineages*", Cancer research 49, 219-225.

Cockburn, J.G., Richardson, D.S., Gujral, T.S., Mulligan, L.M., 2010, "*RET-mediated cell adhesion and migration require multiple integrin subunits*", The Journal of clinical endocrinology and metabolism 95, E342-346.

Constantinescu, R., Constantinescu, A.T., Reichmann, H., Janetzky, B., 2007, "*Neuronal differentiation and long-term culture of the human neuroblastoma line SH-SY5Y*", Journal of neural transmission. Supplementum, 17-28.

Costa, M.W., Lee, S., Furtado, M.B., Xin, L., Sparrow, D.B., Martinez, C.G., Dunwoodie, S.L., Kurtenbach, E., Mohun, T., Rosenthal, N., Harvey, R.P., 2011, "*Complex SUMO-1 regulation of cardiac transcription factor Nkx2-5*", PloS one 6, e24812.

Cragnolini, A.B., **Friedman**, W.J., 2008, "*The function of p75NTR in glia*", Trends Neurosci 31, 99-104.

D'Amelio, M., Cavallucci, V., Cecconi, F., 2010, "*Neuronal caspase-3 signaling: not only cell death*", Cell death and differentiation 17, 1104-1114.

Dai, M., Wang, P., Boyd, A.D., Kostov, G., Athey, B., Jones, E.G., Bunney, W.E., Myers, R.M., Speed, T.P., Akil, H., Watson, S.J., Meng, F., 2005, "*Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data*", Nucleic acids research 33, e175.

Do, C.B., Tung, J.Y., Dorfman, E., Kiefer, A.K., Drabant, E.M., Francke, U., Mountain, J.L., Goldman, S.M., Tanner, C.M., Langston, J.W., Wojcicki, A., Eriksson, N., 2011, "*Web-Based Genome-Wide Association Study Identifies Two Novel Loci and a Substantial Genetic Component for Parkinson's Disease*", PLoS genetics 7, e1002141.

Donnelly, C.R., Gabreski, N.A., Suh, E.B., Chowdhury, M., Pierchala, B.A., 2018, "*Non-canonical Ret signaling augments p75-mediated cell death in developing sympathetic neurons*", The Journal of cell biology 217, 3237-3253.

Doray, B., Salomon, R., Amiel, J., Pelet, A., Touraine, R., Billaud, M., Attie, T., Bachy, B., Munnich, A., Lyonnet, S., 1998, "*Mutation of the RET ligand, neurturin, supports multigenic inheritance in Hirschsprung disease*", Human molecular genetics 7, 1449-1452.

Duval, D., Duval, G., Kedinger, C., Poch, O., Boeuf, H., 2003, "*The 'PINIT' motif, of a newly identified conserved domain of the PIAS protein family, is essential for nuclear retention of PIAS3L*", FEBS letters 554, 111-118.

Eberle, D., Hegarty, B., Bossard, P., Ferre, P., Foufelle, F., 2004, "*SREBP transcription factors: master regulators of lipid homeostasis*", Biochimie 86, 839-848.

Ebmeier, C.C., Taatjes, D.J., 2010, "*Activator-Mediator binding regulates Mediator-cofactor interactions*", Proceedings of the National Academy of Sciences of the United States of America 107, 11283-11288.

El Meskini, R., Crabtree, K.L., Cline, L.B., Mains, R.E., Eipper, B.A., Ronnett, G.V., 2007, "*ATP7A (Menkes protein) functions in axonal targeting and synaptogenesis*", Molecular and cellular neurosciences 34, 409-421.

Emison, E.S., Garcia-Barcelo, M., Grice, E.A., Lantieri, F., Amiel, J., Burzynski, G., Fernandez, R.M., Hao, L., Kashuk, C., West, K., Miao, X., Tam, P.K., Griseri, P., Ceccherini, I., Pelet, A., Jannot, A.S., de Pontual, L., Henrion-Caude, A., Lyonnet, S., Verheij, J.B., Hofstra, R.M., Antinolo, G., Borrego, S., McCallion, A.S., Chakravarti, A., 2010, "*Differential contributions of rare and common, coding and noncoding Ret mutations to multifactorial Hirschsprung disease liability*", American journal of human genetics 87, 60-74.

Emison, E.S., McCallion, A.S., Kashuk, C.S., Bush, R.T., Grice, E., Lin, S., Portnoy, M.E., Cutler, D.J., Green, E.D., Chakravarti, A., 2005, "*A common sex-dependent mutation in a RET enhancer underlies Hirschsprung disease risk*", Nature 434, 857-863.

Encinas, M., Iglesias, M., Liu, Y., Wang, H., Muhaisen, A., Cena, V., Gallego, C., Comella, J.X., 2000, "*Sequential treatment of SH-SY5Y cells with retinoic acid and brain-derived neurotrophic factor gives rise to fully differentiated, neurotrophic factor-dependent, human neuron-like cells*", Journal of neurochemistry 75, 991-1003.

Enomoto, H., Crawford, P.A., Gorodinsky, A., Heuckeroth, R.O., Johnson, E.M., Jr., Milbrandt, J., 2001, "*RET signaling is essential for migration, axonal growth and axon guidance of developing sympathetic neurons*", Development (Cambridge, England) 128, 3963-3974.

Esposito, C.L., D'Alessio, A., de Franciscis, V., Cerchia, L., 2008, "*A Cross-Talk between TrkB and Ret Tyrosine Kinases Receptors Mediates Neuroblastoma Cells Differentiation*", PloS one 3.

Estruch, S.B., Graham, S.A., Deriziotis, P., Fisher, S.E., 2016, "*The language-related transcription factor FOXP2 is post-translationally modified with small ubiquitin-like modifiers*", Scientific reports 6, 20911.

Fattahi, F., Steinbeck, J.A., Kriks, S., Tchieu, J., Zimmer, B., Kishinevsky, S., Zeltner, N., Mica, Y., El-Nachef, W., Zhao, H., de Stanchina, E., Gershon, M.D., Grikscheit, T.C., Chen, S., Studer, L., 2016, "*Deriving human ENS lineages for cell therapy and drug discovery in Hirschsprung disease*", Nature 531, 105-109.

Faure, C., Thapa, N., Di Lorenzo, C., Heanue, T.A., Burns, A.J., Beckett E.A., Young, H.M., Bornstein, J.C., Jadcherla, S.R., Righini-Grunder, F., Preidis, G.A., Chumpitazi, B.P., Shulman, R.J., van Tilburg, M.A.L., Turco, R., Staiano, A., Saliakellis, E., Rybak, A., Borrelli, O., Yacob, D., Di Lorenzo, C., Zar-Kessler, C., Belkind-Gerson, J., Rommel, N., Chiou, E., Rosen, RL., Biassoni, L., Easty, M., Arbizu, R.A., Rodriguez, L., Chelimsky, G., Chelimsky, T.C., Kapur, R.P., Gisser, M., Gariepy, C.E., Nibali, R.C., Duncan, D.R., Aspirot, A., Gomez, R., Furtonato, J.E, Nurko, S., Vandenplas, Y., Sung, V., James-Roberts, I., Guiraldes, E., Roessler, J.L., Rosen, M., Saps, M., Mokha, J.S., Hyams, J.S., Sood, M.R., Kovacic, K., Li, B.U.K., Alioto, A., Koppen, I.J.N., Benninga, M.A., Har, A.F., Croffie, J.M.B., Hyman, P.E., Arrouk, R., Teich, S., Vlieger, A.M., Hoota, R., Natarajan, D., Goulet, O., Irtan, S., 2017, "*Pediatric Neurogastroenterology: Gastrointestinal Motility and Functional Disorders in Children*", 2nd ed. Springer International Publishing.

Fromer, M., **Purcell**, S.M., 2014, "*Using XHMM Software to Detect Copy Number Variation in Whole-Exome Sequencing Data*", Curr Protoc Hum Genet 81, 7.23.21-21.

Fu, M., Sato, Y., Lyons-Warren, A., Zhang, B., Kane, M.A., Napoli, J.L., Heuckeroth, R.O., 2010, "*Vitamin A facilitates enteric nervous system precursor migration by reducing Pten accumulation*", Development (Cambridge, England) 137, 631-640.

Fu, M., Tam, P.K., Sham, M.H., Lui, V.C., 2004, "*Embryonic development of the ganglion plexuses and the concentric layer structure of human gut: a topographical study*", Anatomy and embryology 208, 33-41.

Furness, J.B., Stebbing, M.J., 2018, "*The first brain: Species comparisons and evolutionary implications for the enteric and central nervous systems*", Neurogastroenterology and motility : the official journal of the European Gastrointestinal Motility Society 30.

Gammill, L.S., Bronner-Fraser, M., 2003, "*Neural crest specification: migrating into genomics*", Nature reviews. Neuroscience 4, 795-805.

Geillon, F., Gondcaille, C., Charbonnier, S., Van Roermund, C.W., Lopez, T.E., Dias, A.M., Pais de Barros, J.P., Arnould, C., Wanders, R.J., Trompier, D., Savary, S., 2014, "*Structure-function analysis of peroxisomal ATP-binding cassette transporters using chimeric dimers*", The Journal of biological chemistry 289, 24511-24520.

Gilbert, S.F., 2014, "*Developmental biology*", 10th ed. Sinauer, A. D., Sunderland.

Goldie, B.J., Barnett, M.M., Cairns, M.J., 2014, "*BDNF and the maturation of posttranscriptional regulatory networks in human SH-SY5Y neuroblast differentiation*", Frontiers in cellular neuroscience 8, 325.

Goldstein, A.M., Hofstra, R.M., Burns, A.J., 2013, "*Building a brain in the gut: development of the enteric nervous system*", Clin Genet 83, 307-316.

Gourdon, P., Sitsel, O., Lykkegaard Karlsen, J., Birk Moller, L., Nissen, P., 2012, "*Structural models of the human copper P-type ATPases ATP7A and ATP7B*", Biol Chem 393, 205-216.

Graham, H.K., Maina, I., Goldstein, A.M., Nagy, N., 2017, "*Intestinal smooth muscle is required for patterning the enteric nervous system*", Journal of anatomy 230, 567-574.

Griffiths, B., Lewis, C.A., Bensaad, K., Ros, S., Zhang, Q., Ferber, E.C., Konisti, S., Peck, B., Miess, H., East, P., Wakelam, M., Harris, A.L., Schulze, A., 2013, "*Sterol regulatory element binding protein-dependent regulation of lipid synthesis supports cell survival and tumor growth*", Cancer & metabolism 1, 3.

Grosse, A.S., Pressprich, M.F., Curley, L.B., Hamilton, K.L., Margolis, B., Hildebrand, J.D., Gumucio, D.L., 2011, "*Cell dynamics in fetal intestinal epithelium: implications for intestinal growth and morphogenesis*", Development (Cambridge, England) 138, 4423-4432.

Gui, H., Schriemer, D., Cheng, W.W., Chauhan, R.K., Antinolo, G., Berrios, C., Bleda, M., Brooks, A.S., Brouwer, R.W., Burns, A.J., Cherny, S.S., Dopazo, J., Eggen, B.J., Griseri, P., Jalloh, B., Le, T.L., Lui, V.C., Luzon-Toro, B., Matera, I., Ngan, E.S., Pelet, A., Ruiz-Ferrer, M., Sham, P.C., Shepherd, I.T., So, M.T., Sribudiani, Y., Tang, C.S., van den Hout, M.C., van der Linde, H.C., van Ham, T.J., van, I.W.F., Verheij, J.B., Amiel, J., Borrego, S., Ceccherini, I., Chakravarti, A., Lyonnet, S., Tam, P.K., Garcia-Barcelo, M.M., Hofstra, R.M., 2017, "*Whole exome sequencing coupled with unbiased functional analysis reveals new Hirschsprung disease genes*", Genome biology 18, 48.

Guo, M.H., Plummer, L., Chan, Y.M., Hirschhorn, J.N., Lippincott, M.F., 2018, "*Burden Testing of Rare Variants Identified through Exome Sequencing via Publicly Available Control Data*", American journal of human genetics 103, 522-534.

Hao, M.M., Bornstein, J.C., Vanden Berghe, P., Lomax, A.E., Young, H.M., Foong, J.P., 2013, "*The emergence of neural activity and its role in the development of the enteric nervous system*", Developmental biology 382, 365-374.

Hao, M.M., **Young**, H.M., 2009, "*Development of enteric neuron diversity*", Journal of cellular and molecular medicine 13, 1193-1210.

Heanue, T.A., **Pachnis**, V., 2007, "*Enteric nervous system development and Hirschsprung's disease: advances in genetic and stem cell studies*", Nature reviews. Neuroscience 8, 466-479.

Heanue, T.A., **Pachnis**, V., 2006, "*Expression profiling the developing mammalian enteric nervous system identifies marker and candidate Hirschsprung disease genes*", Proceedings of the National Academy of Sciences of the United States of America 103, 6919-6924.

Heuckeroth, R.O., 2015, "*Hirschsprung's disease, Down syndrome, and missing heritability: too much collagen slows migration*", The Journal of clinical investigation 125, 4323-4326.

Heuckeroth, R.O., Enomoto, H., Grider, J.R., Golden, J.P., Hanke, J.A., Jackman, A., Molliver, D.C., Bardgett, M.E., Snider, W.D., Johnson, E.M., Jr., Milbrandt, J., 1999, "*Gene targeting reveals a critical role for neurturin in the development and maintenance of enteric, sensory, and parasympathetic neurons*", Neuron 22, 253-263.

Heuckeroth, R.O., Schäfer, K.H., 2016, "*Gene-environment interactions and the enteric nervous system: Neural plasticity and Hirschsprung disease prevention*", Developmental biology 417, 188-197.

Hirschsprung, H., 1888, "*Stuhlträgheit Neugeborener in Folge von Dilatation und Hypertrophie des Colons*.

Hofstra, R.M., Valdenaire, O., Arch, E., Osinga, J., Kroes, H., Loffler, B.M., Hamosh, A., Meijers, C., Buys, C.H., 1999, "*A loss-of-function mutation in the endothelin-converting enzyme 1 (ECE-1) associated with Hirschsprung disease, cardiac defects, and autonomic dysfunction*", American journal of human genetics 64, 304-308.

Höftberger, R., Kunze, M., Weinhofer, I., Aboul-Enein, F., Voigtlander, T., Oezen, I., Amann, G., Bernheimer, H., Budka, H., Berger, J., 2007, "*Distribution and cellular localization of adrenoleukodystrophy protein in human tissues: implications for X-linked adrenoleukodystrophy*", Neurobiology of disease 28, 165-174.

Horton, J.D., Goldstein, J.L., Brown, M.S., 2002, "*SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver*", The Journal of clinical investigation 109, 1125-1131.

Ibáñez, C.F., 2013, "*Structure and physiology of the RET receptor tyrosine kinase*", Cold Spring Harbor perspectives in biology 5.

Im, S.-S., Hammond, L.E., Yousef, L., Nugas-Selby, C., Shin, D.-J., Seo, Y.-K., Fong, L.G., Young, S.G., Osborne, T.F., 2009, "*Sterol Regulatory Element Binding Protein 1a Regulates Hepatic Fatty Acid Partitioning by Activating Acetyl Coenzyme A Carboxylase 2*", 29, 4864-4872.

Jang, J., Kang, H.C., Kim, H.S., Kim, J.Y., Huh, Y.J., Kim, D.S., Yoo, J.E., Lee, J.A., Lim, B., Lee, J., Yoon, T.M., Park, I.H., Hwang, D.Y., Daley, G.Q., Kim, D.W., 2011, "*Induced pluripotent stem cell models from X-linked adrenoleukodystrophy patients*", Annals of neurology 70, 402-409.

Jeong, S., Liang, G., Sharma, S., Lin, J.C., Choi, S.H., Han, H., Yoo, C.B., Egger, G., Yang, A.S., Jones, P.A., 2009, "*Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA*", Molecular and cellular biology 29, 5366-5376.

Jiang, M., Stanke, J., Lahti, J.M., 2011a, "*The connections between neural crest development and neuroblastoma*", Current topics in developmental biology 94, 77-127.

Jiang, Q., Arnold, S., Heanue, T., Kilambi, K.P., Doan, B., Kapoor, A., Ling, A.Y., Sosa, M.X., Guy, M., Jiang, Q., Burzynski, G., West, K., Bessling, S., Griseri, P., Amiel, J., Fernandez, R.M., Verheij, J.B., Hofstra, R.M., Borrego, S., Lyonnet, S., Ceccherini, I., Gray, J.J., Pachnis, V., McCallion, A.S., Chakravarti, A., 2015, "*Functional loss of semaphorin 3C and/or semaphorin 3D and their epistatic interaction with ret are critical to Hirschsprung disease liability*", American journal of human genetics 96, 581-596.

Jiang, Q., Ho, Y.Y., Hao, L., Nichols Berrios, C., Chakravarti, A., 2011b, "*Copy number variants in candidate genes are genetic modifiers of Hirschsprung disease*", PloS one 6, e21219.

Johnsen, J.I., Dyberg, C., Wickstrom, M., 2019, "*Neuroblastoma-A Neural Crest Derived Embryonal Malignancy*", Frontiers in molecular neuroscience 12, 9.

Kaler, S.G., 2011, "*ATP7A-related copper transport diseases-emerging concepts and future trends*", Nature reviews. Neurology 7, 15-29.

Kapoor, A., Jiang, Q., Chatterjee, S., Chakraborty, P., Sosa, M.X., Berrios, C., Chakravarti, A., 2015, "*Population variation in total genetic risk of Hirschsprung disease from common RET, SEMA3 and NRG1 susceptibility polymorphisms*", Human molecular genetics 24, 2997-3003.

Kawaguchi, K., Morita, M., 2016, "*ABC Transporter Subfamily D: Distinct Differences in Behavior between ABCD1-3 and ABCD4 in Subcellular Localization, Function, and Human Disease*", BioMed research international 2016, 6786245.

Kemp, S., Theodoulou, F.L., Wanders, R.J., 2011, "*Mammalian peroxisomal ABC transporters: from endogenous substrates to pathology and clinical significance*", Br J Pharmacol 164, 1753-1766.

Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., Shendure, J., 2014, "*A general framework for estimating the relative pathogenicity of human genetic variants*", Nature genetics 46, 310-315.

Korecka, J.A., van Kesteren, R.E., Blaas, E., Spitzer, S.O., Kamstra, J.H., Smit, A.B., Swaab, D.F., Verhaagen, J., Bossers, K., 2013, "*Phenotypic characterization of retinoic acid differentiated SH-SY5Y cells by transcriptional profiling*", PloS one 8, e63862.

Kosicki, M., Tomberg, K., Bradley, A., 2018, "*Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements*", Nature Biotechnology 36, 765.

Kotaja, N., Karvonen, U., Janne, O.A., Palvimo, J.J., 2002, "*PIAS proteins modulate transcription factors by functioning as SUMO-1 ligases*", Molecular and cellular biology 22, 5222-5234.

Kovalevich, J., Langford, D., 2013, "*Considerations for the use of SH-SY5Y neuroblastoma cells in neurobiology*", Methods in molecular biology (Clifton, N.J.) 1078, 9-21.

Kruger, G.M., Mosher, J.T., Bixby, S., Joseph, N., Iwashita, T., Morrison, S.J., 2002, "*Neural crest stem cells persist in the adult gut but undergo changes in self-renewal, neuronal subtype potential, and factor responsiveness*", Neuron 35, 657-669.

Kurokawa, K., Kawai, K., Hashimoto, M., Ito, Y., Takahashi, M., 2003, "*Cell signalling and gene expression mediated by RET tyrosine kinase*", Journal of internal medicine 253, 627-633.

Lai, F.P., Lau, S.T., Wong, J.K., Gui, H., Wang, R.X., Zhou, T., Lai, W.H., Tse, H.F., Tam, P.K., Garcia-Barcelo, M.M., Ngan, E.S., 2017, "*Correction of Hirschsprung-Associated Mutations in Human Induced Pluripotent Stem Cells Via Clustered Regularly Interspaced Short Palindromic Repeats/Cas9, Restores Neural Crest Cell Function*", Gastroenterology 153, 139-153.e138.

Lake, J.I., **Heuckeroth**, R.O., 2013, "*Enteric nervous system development: migration, differentiation, and disease*", American journal of physiology. Gastrointestinal and liver physiology 305, G1-24.

Lalonde, S., Stone, O.A., Lessard, S., Lavertu, A., Desjardins, J., Beaudoin, M., Rivas, M., Stainier, D.Y.R., Lettre, G., 2017, "*Frameshift indels introduced by genome editing can lead to in-frame exon skipping*", PloS one 12, e0178700.

Langmann, T., Mauerer, R., Zahn, A., Moehle, C., Probst, M., Stremmel, W., Schmitz, G., 2003, "*Real-time reverse transcription-PCR expression profiling of the complete human ATP-binding cassette transporter superfamily in various tissues*", Clinical chemistry 49, 230-238.

Lee, A., Asahina, K., Okamoto, T., Kawaguchi, K., Kostsin, D.G., Kashiwayama, Y., Takanashi, K., Yazaki, K., Imanaka, T., Morita, M., 2014a, "*Role of NH2-terminal hydrophobic motif in the subcellular localization of ATP-binding cassette protein subfamily D: common features in eukaryotic organisms*", Biochemical and biophysical research communications 453, 612-618.

Lee, J.H., Lee, G.Y., Jang, H., Choe, S.S., Koo, S.H., Kim, J.B., 2014b, "*Ring finger protein20 regulates hepatic lipid metabolism through protein kinase A-dependent sterol regulatory element binding protein1c degradation*", Hepatology 60, 844-857.

Lee, S., Abecasis, G.R., Boehnke, M., Lin, X., 2014c, "*Rare-variant association analysis: study designs and statistical tests*", American journal of human genetics 95, 5-23.

Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., Tukiainen, T., Birnbaum, D.P., Kosmicki, J.A., Duncan, L.E., Estrada, K., Zhao, F., Zou, J., Pierce-Hoffman, E., Berghout, J., Cooper, D.N., Deflaux, N., DePristo, M., Do, R., Flannick, J., Fromer, M., Gauthier, L., Goldstein, J., Gupta, N., Howrigan, D., Kiezun, A., Kurki, M.I., Moonshine, A.L., Natarajan, P., Orozco, L., Peloso, G.M., Poplin, R., Rivas, M.A., Ruano-Rubio, V., Rose, S.A., Ruderfer, D.M., Shakir, K., Stenson, P.D., Stevens, C., Thomas, B.P., Tiao, G., Tusie-Luna, M.T., Weisburd, B., Won, H.H., Yu, D., Altshuler, D.M., Ardissino, D., Boehnke, M., Danesh, J., Donnelly, S., Elosua, R., Florez, J.C., Gabriel, S.B., Getz, G., Glatt, S.J., Hultman, C.M., Kathiresan, S., Laakso, M., McCarroll, S., McCarthy, M.I., McGovern, D., McPherson, R., Neale, B.M., Palotie, A., Purcell, S.M., Saleheen, D., Scharf, J.M., Sklar, P., Sullivan, P.F., Tuomilehto, J., Tsuang, M.T., Watkins, H.C., Wilson, J.G., Daly, M.J., MacArthur, D.G., 2016, "*Analysis of protein-coding genetic variation in 60,706 humans*", Nature 536, 285-291.

Lopez-Carballo, G., Moreno, L., Masia, S., Perez, P., Barettino, D., 2002, "*Activation of the phosphatidylinositol 3-kinase/Akt signaling pathway by retinoic acid is required for neural differentiation of SH-SY5Y human neuroblastoma cells*", The Journal of biological chemistry 277, 25297-25304.

Luzón-Toro, B., Espino-Paisán, L., Fernández, R.M., Martín-Sánchez, M., Antiñolo, G., Borrego, S., 2015a, "*Next-generation-based targeted sequencing as an efficient tool for the study of the genetic background in Hirschsprung patients*", BMC medical genetics 16, 89.

Luzón-Toro, B., Gui, H., Ruiz-Ferrer, M., Sze-Man Tang, C., Fernandez, R.M., Sham, P.C., Torroglosa, A., Kwong-Hang Tam, P., Espino-Paisan, L., Cherny, S.S., Bleda, M., Enguix-Riego Mdel, V., Dopazo, J., Antinolo, G., Garcia-Barcelo, M.M., Borrego, S., 2015b, "*Exome sequencing reveals a high genetic heterogeneity on familial Hirschsprung disease*", Scientific reports 5, 16473.

Martin, S., Wilkinson, K.A., Nishimune, A., Henley, J.M., 2007, "*Emerging extranuclear roles of protein SUMOylation in neuronal function and dysfunction*", Nature reviews. Neuroscience 8, 948-959.

Martins-Taylor, K., Schroeder, D.I., LaSalle, J.M., Lalande, M., Xu, R.H., 2012, "*Role of DNMT3B in the regulation of early neural and neural crest specifiers*", Epigenetics 7, 71-82.

Mayer, E.A., **Tillisch**, K., 2011, "*The brain-gut axis in abdominal pain syndromes*", Annual review of medicine 62, 381-396.

Mayor, R., **Theveneau**, E., 2013, "*The neural crest*", Development (Cambridge, England) 140, 2247-2251.

McCallion, A.S., Stames, E., Conlon, R.A., Chakravarti, A., 2003, "*Phenotype variation in two-locus mouse models of Hirschsprung disease: tissue-specific interaction between Ret and Ednrb*", Proceedings of the National Academy of Sciences of the United States of America 100, 1826-1831.

Meeker, R.B., **Williams**, K.S., 2015, "*The p75 neurotrophin receptor: at the crossroad of neural repair and death*", Neural regeneration research 10, 721-725.

Meka, D.P., Muller-Rischart, A.K., Nidadavolu, P., Mohammadi, B., Motori, E., Ponna, S.K., Aboutalebi, H., Bassal, M., Annamneedi, A., Finckh, B., Miesbauer, M., Rotermund, N., Lohr, C., Tatzelt, J., Winklhofer, K.F., Kramer, E.R., 2015, "*Parkin cooperates with GDNF/RET signaling to prevent dopaminergic neuron degeneration*", The Journal of clinical investigation 125, 1873-1885.

Memic, F., Knoflach, V., Morarach, K., Sadler, R., Laranjeira, C., Hjerling-Leffler, J., Sundstrom, E., Pachnis, V., Marklund, U., 2018, "*Transcription and Signaling Regulators in Developing Neuronal Subtypes of Mouse and Human Enteric Nervous System*", Gastroenterology 154, 624-636.

Mendelsohn, B.A., Yin, C., Johnson, S.L., Wilm, T.P., Solnica-Krezel, L., Gitlin, J.D., 2006, "*Atp7a determines a hierarchy of copper metabolism essential for notochord development*", Cell metabolism 4, 155-162.

Meraldi, P., 2019, "*Bub1-the zombie protein that CRISPR cannot kill*", The EMBO journal 38.

Messi, E., Florian, M.C., Caccia, C., Zanisi, M., Maggi, R., 2008, "*Retinoic acid reduces human neuroblastoma cell migration and invasiveness: effects on DCX, LIS1, neurofilaments-68 and vimentin expression*", BMC cancer 8, 30.

Metzger, M., Caldwell, C., Barlow, A.J., Burns, A.J., Thapar, N., 2009, "*Enteric nervous system stem cells derived from human gut mucosa for the treatment of aganglionic gut disorders*", Gastroenterology 136, 2214-2225.e2211-2213.

Miller, I., Min, M., Yang, C., Tian, C., Gookin, S., Carter, D., Spencer, S.L., 2018, "*Ki67 is a Graded Rather than a Binary Marker of Proliferation versus Quiescence*", Cell reports 24, 1105-1112.e1105.

Monnerie, H., Romer, M., Jensen, B.K., Millar, J.S., Jordan-Sciutto, K.L., Kim, S.F., Grinspan, J.B., 2017, "*Reduced sterol regulatory element-binding protein (SREBP) processing through site-1 protease (S1P) inhibition alters oligodendrocyte differentiation in vitro*", Journal of neurochemistry 140, 53-67.

Montague, T.G., Cruz, J.M., Gagnon, J.A., Church, G.M., Valen, E., 2014, "*CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing*", Nucleic acids research 42, W401-407.

Morita, M., **Imanaka**, T., 2012, "*Peroxisomal ABC transporters: structure, function and role in disease*", Biochimica et biophysica acta 1822, 1387-1396.

Morvay, P.L., Baes, M., Van Veldhoven, P.P., 2017, "*Differential activities of peroxisomes along the mouse intestinal epithelium*", Cell Biochem Funct 35, 144-155.

Mou, H., Smith, J.L., Peng, L., Yin, H., Moore, J., Zhang, X.O., Song, C.Q., Sheel, A., Wu, Q., Ozata, D.M., Li, Y., Anderson, D.G., Emerson, C.P., Sontheimer, E.J., Moore, M.J., Weng, Z., Xue, W., 2017, "*CRISPR/Cas9-mediated genome editing induces exon skipping by alternative splicing or exon deletion*", Genome biology 18, 108.

Murillo, J.R., Goto-Silva, L., Sanchez, A., Nogueira, F.C.S., Domont, G.B., Junqueira, M., 2017, "*Quantitative proteomic analysis identifies proteins and pathways related to neuronal development in differentiated SH-SY5Y neuroblastoma cells*", EuPA open proteomics 16, 1-11.

Mutch, D.M., Anderle, P., Fiaux, M., Mansourian, R., Vidal, K., Wahli, W., Williamson, G., Roberts, M.A., 2004, "*Regional variations in ABC transporter expression along the mouse intestinal tract*", Physiological genomics 17, 11-20.

Nagy, N., Barad, C., Graham, H.K., Hotta, R., Cheng, L.S., Fejszak, N., Goldstein, A.M., 2016, "*Sonic hedgehog controls enteric nervous system development by patterning the extracellular matrix*", Development (Cambridge, England) 143, 264-275.

Obermayr, F., Hotta, R., Enomoto, H., Young, H.M., 2013, "*Development and developmental disorders of the enteric nervous system*", Nature reviews. Gastroenterology & hepatology 10, 43-57.

Obermayr, F., **Seitz**, G., 2018, "*Recent developments in cell-based ENS regeneration – a short review*", iss 3, 93.

Oberoi, R., Bogalle, E.P., Matthes, L.A., Schuett, H., Koch, A.K., Grote, K., Schieffer, B., Schuett, J., Luchtefeld, M., 2015, "*Lipocalin (LCN) 2 Mediates Pro-Atherosclerotic Processes and Is Elevated in Patients with Coronary Artery Disease*", PloS one 10, e0137924.

Ong, W.Y., Hu, C.Y., Soh, Y.P., Lim, T.M., Pentchev, P.G., Patel, S.C., 2000, "*Neuronal localization of sterol regulatory element binding protein-1 in the rodent and primate brain: a light and electron microscopic immunocytochemical study*", Neuroscience 97, 143-153.

Palsson, O.S., Whitehead, W.E., van Tilburg, M.A., Chang, L., Chey, W., Crowell, M.D., Keefer, L., Lembo, A.J., Parkman, H.P., Rao, S.S., Sperber, A., Spiegel, B., Tack, J., Vanner, S., Walker, L.S., Whorwell, P., Yang, Y., 2016, "*Rome IV Diagnostic Questionnaires and Tables for Investigators and Clinicians*", Gastroenterology.

Parisi, M.A., 2015, "*Hirschsprung Disease Overview*", GeneReviews, Adam, M.P., Ardinger, H.H., Paggon, R. A., Wallace, S.E. (Eds), University of Seattle, ISSN: 2372-0697

Parisi, M.A., **Kapur**, R.P., 2000, "*Genetics of Hirschsprung disease*", Current opinion in pediatrics 12, 610-617.

Pfundt, R., Del Rosario, M., Vissers, L., Kwint, M.P., Janssen, I.M., de Leeuw, N., Yntema, H.G., Nelen, M.R., Lugtenberg, D., Kamsteeg, E.J., Wieskamp, N., Stegmann, A.P.A., Stevens, S.J.C., Rodenburg, R.J.T., Simons, A., Mensenkamp, A.R., Rinne, T., Gilissen, C., Scheffer, H., Veltman, J.A.P.D., Hehir-Kwa, J.Y., 2017, "*Detection of clinically relevant copy-number variants by exome sequencing in a large cohort of genetic disorders*", Genetics in medicine : official journal of the American College of Medical Genetics 19, 667-675.

Pingault, V., Bondurand, N., Kuhlbrodt, K., Goerich, D.E., Prehu, M.O., Puliti, A., Herbarth, B., Hermans-Borgmeyer, I., Legius, E., Matthijs, G., Amiel, J., Lyonnet, S., Ceccherini, I., Romeo, G., Smith, J.C., Read, A.P., Wegner, M., Goossens, M., 1998, "*SOX10 mutations in patients with Waardenburg-Hirschsprung disease*", Nature genetics 18, 171-173.

Porstmann, T., Santos, C.R., Griffiths, B., Cully, M., Wu, M., Leevers, S., Griffiths, J.R., Chung, Y.L., Schulze, A., 2008, "*SREBP activity is regulated by mTORC1 and contributes to Akt-dependent cell growth*", Cell metabolism 8, 224-236.

Puig, I., Champeval, D., De Santa Barbara, P., Jaubert, F., Lyonnet, S., Larue, L., 2009, "*Deletion of Pten in the mouse enteric nervous system induces ganglioneuromatosis and mimics intestinal pseudoobstruction*", The Journal of clinical investigation 119, 3586-3596.

Qin, Z., Konaniah, E.S., Neltner, B., Nemenoff, R.A., Hui, D.Y., Weintraub, N.L., 2010, "*Participation of ATP7A in macrophage mediated oxidation of LDL*", Journal of lipid research 51, 1471-1477.

Raas, Q., Gondcaille, C., Hamon, Y., Leoni, V., Caccia, C., Menetrier, F., Lizard, G., Trompier, D., Savary, S., 2019, "*CRISPR/Cas9-mediated knockout of Abcd1 and Abcd2 genes in BV-2 cells: novel microglial models for X-linked Adrenoleukodystrophy*", Biochimica et biophysica acta. Molecular and cell biology of lipids 1864, 704-714.

Raghavan, S., **Bitar**, K.N., 2014, "*The influence of extracellular matrix composition on the differentiation of neuronal subtypes in tissue engineered innervated intestinal smooth muscle sheets*", Biomaterials 35, 7429-7440.

Rao, M., & **Gershon**, M. D. , 2016, "*The bowel and beyond: the enteric nervous system in neurological disorders.*", Gastroenterology & Hepatology Nature Reviews. 13, 517–528.

Rao, M., **Gershon**, M.D., 2018, "*Enteric nervous system development: what could possibly go wrong?*", Nature reviews. Neuroscience 19, 552-565.

Reber, S., Mechtersheimer, J., Nasif, S., Benitez, J.A., Colombo, M., Domanski, M., Jutzi, D., Hedlund, E., Ruepp, M.D., 2018, "*CRISPR-Trap: a clean approach for the generation of gene knockouts and gene replacements in human cells*", Molecular biology of the cell 29, 75-83.

Rolland, T., Tasan, M., Charloteaux, B., Pevzner, S.J., Zhong, Q., Sahni, N., Yi, S., Lemmens, I., Fontanillo, C., Mosca, R., Kamburov, A., Ghiassian, S.D., Yang, X., Ghamsari, L., Balcha, D., Begg, B.E., Braun, P., Brehme, M., Broly, M.P., Carvunis, A.R., Convery-Zupan, D., Corominas, R., Coulombe-Huntington, J., Dann, E., Dreze, M., Dricot, A., Fan, C., Franzosa, E., Gebreab, F., Gutierrez, B.J., Hardy, M.F., Jin, M., Kang, S., Kiros, R., Lin, G.N., Luck, K., MacWilliams, A., Menche, J., Murray, R.R., Palagi, A., Poulin, M.M., Rambout, X., Rasla, J., Reichert, P., Romero, V., Ruyssinck, E., Sahalie, J.M., Scholz, A., Shah, A.A., Sharma, A., Shen, Y., Spirohn, K., Tam, S., Tejeda, A.O., Trigg, S.A., Twizere, J.C., Vega, K., Walsh, J., Cusick, M.E., Xia, Y., Barabasi, A.L., Iakoucheva, L.M., Aloy, P., De Las Rivas, J., Tavernier, J., Calderwood, M.A., Hill, D.E., Hao, T., Roth, F.P., Vidal, M., 2014, "*A proteome-scale map of the human interactome network*", Cell 159, 1212-1226.

Rollo, B.N., Zhang, D., Stamp, L.A., Menheniott, T.R., Stathopoulos, L., Denham, M., Dottori, M., King, S.K., Hutson, J.M., Newgreen, D.F., 2016, "*Enteric Neural Cells From Hirschsprung Disease Patients Form Ganglia in Autologous Aneuronal Colon*", Cellular and molecular gastroenterology and hepatology 2, 92-109.

Ross, R.A., Spengler, B.A., Biedler, J.L., 1983, "*Coordinate morphological and biochemical interconversion of human neuroblastoma cells*", J Natl Cancer Inst 71, 741-747.

Rothman, T.P., Tennyson, V.M., Gershon, M.D., 1986, "*Colonization of the bowel by the precursors of enteric glia: studies of normal and congenitally aganglionic mutant mice*", J Comp Neurol 252, 493-506.

Rott, R., Szargel, R., Shani, V., Hamza, H., Savyon, M., Abd Elghani, F., Bandopadhyay, R., Engelender, S., 2017, "*SUMOylation and ubiquitination reciprocally regulate α-synuclein degradation and pathological aggregation*", Proceedings of the National Academy of Sciences of the United States of America 114, 13176-13181.

Sasselli, V., Pachnis, V., Burns, A.J., 2012, "*The enteric nervous system*", Developmental biology 366, 64-73.

Schmitteckert, S., Mederer, T., Roth, R., Gunther, P., Holland-Cunz, S., Metzger, M., Samstag, Y., Schroder-Braunstein, J., Wabnitz, G., Kurzhals, S., Scheuerer, J., Beretta, C.A., Lasitschka, F., Rappold, G.A., Romero, P., Niesler, B., 2019, "*Postnatal human enteric neurospheres show a remarkable molecular complexity*", Neurogastroenterology and motility : the official journal of the European Gastrointestinal Motility Society, e13674.

Shalizi, A., Bilimoria, P.M., Stegmuller, J., Gaudilliere, B., Yang, Y., Shuai, K., Bonni, A., 2007, "*PIASx is a MEF2 SUMO E3 ligase that promotes postsynaptic dendritic morphogenesis*", The Journal of neuroscience : the official journal of the Society for Neuroscience 27, 10037-10046.

Shao, W., Espenshade, P.J., 2014, "*Sterol regulatory element-binding protein (SREBP) cleavage regulates Golgi-to-endoplasmic reticulum recycling of SREBP cleavage-activating protein (SCAP)*", The Journal of biological chemistry 289, 7547-7557.

Sharan, A., Zhu, H., Xie, H., Li, H., Tang, J., Tang, W., Zhang, H., Xia, Y., 2015, "*Down-regulation of miR-206 is associated with Hirschsprung disease and suppresses cell migration and proliferation in cell models*", Scientific reports 5, 9302.

Shen, Q., Zhang, H., Su, Y., Wen, Z., Zhu, Z., Chen, G., Peng, L., Du, C., Xie, H., Li, H., Lv, X., Lu, C., Xia, Y., Tang, W., 2018, "*Identification of two novel PCDHA9 mutations associated with Hirschsprung's disease*", Gene 658, 96-104.

Shepherd, I., **Eisen**, J., 2011, "*Development of the zebrafish enteric nervous system*", Methods in cell biology 101, 143-160.

Shimomura, I., Shimano, H., Horton, J.D., Goldstein, J.L., Brown, M.S., 1997, "*Differential expression of exons 1a and 1c in mRNAs for sterol regulatory element binding protein-1 in human and mouse organs and cultured cells*", The Journal of clinical investigation 99, 838-845.

Shuai, K., 2000, "*Modulation of STAT signaling by STAT-interacting proteins*", Oncogene 19, 2638-2644.

Simões-Costa, M., **Bronner**, M.E., 2015, "*Establishing neural crest identity: a gene regulatory recipe*", Development (Cambridge, England) 142, 242-257.

Singh, S., Shariff, A., Roy, T., Das, T., Rani, N., 2015, "*Development of myenteric plexus in human foetuses: a quantitative study*", Anatomy & cell biology 48, 124-129.

Skjørringe, T., Amstrup Pedersen, P., Salling Thorborg, S., Nissen, P., Gourdon, P., Birk Møller, L., 2017, "*Characterization of ATP7A missense mutants suggests a correlation between intracellular trafficking and severity of Menkes disease*", Scientific reports 7, 757.

Smith, J.L., Mou, H., Xue, W., 2018, "*Understanding and repurposing CRISPR-mediated alternative splicing*", Genome biology 19, 184.

So, M.T., Leon, T.Y., Cheng, G., Tang, C.S., Miao, X.P., Cornes, B.K., Diem, N.N., Cui, L., Ngan, E.S., Lui, V.C., Wu, X.Z., Wang, B., Wang, H., Yuan, Z.W., Huang, L.M., Li, L., Xia, H., Zhu, D., Liu, J., Nguyen, T.L., Chan, I.H., Chung, P.H., Liu, X.L., Zhang, R., Wong, K.K., Sham, P.C., Cherny, S.S., Tam, P.K., Garcia-Barcelo, M.M., 2011, "*RET mutational spectrum in Hirschsprung disease: evaluation of 601 Chinese patients*", PloS one 6, e28986.

Stemmer, M., Thumberger, T., Del Sol Keyer, M., Wittbrodt, J., Mateo, J.L., 2015, "*CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool*", PloS one 10, e0124633.

Strachan, L.R., Stevenson, T.J., Freshner, B., Keefe, M.D., Miranda Bowles, D., Bonkowsky, J.L., 2017, "*A zebrafish model of X-linked adrenoleukodystrophy recapitulates key disease features and demonstrates a developmental requirement for abcd1 in oligodendrocyte patterning and myelination*", Human molecular genetics 26, 3600-3614.

Tang, C.S., Cheng, G., So, M.T., Yip, B.H., Miao, X.P., Wong, E.H., Ngan, E.S., Lui, V.C., Song, Y.Q., Chan, D., Cheung, K., Yuan, Z.W., Lei, L., Chung, P.H., Liu, X.L., Wong, K.K., Marshall, C.R., Scherer, S.W., Cherny, S.S., Sham, P.C., Tam, P.K., Garcia-Barcelo, M.M., 2012, "*Genome-wide copy number analysis uncovers a new HSCR gene: NRG3*", PLoS genetics 8, e1002687.

Taraviras, S., Marcos-Gutierrez, C.V., Durbec, P., Jani, H., Grigoriou, M., Sukumaran, M., Wang, L.C., Hynes, M., Raisman, G., Pachnis, V., 1999, "*Signalling by the RET receptor tyrosine kinase and its role in the development of the mammalian enteric nervous system*", Development (Cambridge, England) 126, 2785-2797.

Telianidis, J., Hung, Y.H., Materia, S., Fontaine, S.L., 2013, "*Role of the P-Type ATPases, ATP7A and ATP7B in brain copper homeostasis*", Frontiers in aging neuroscience 5, 44.

Tilghman, J.M., Ling, A.Y., Turner, T.N., Sosa, M.X., Krumm, N., Chatterjee, S., Kapoor, A., Coe, B.P., Nguyen, K.H., Gupta, N., Gabriel, S., Eichler, E.E., Berrios, C., Chakravarti, A., 2019, "*Molecular Genetic Anatomy and Risk Profile of Hirschsprung's Disease*", N Engl J Med 380, 1421-1432.

Torroglosa, A., Enguix-Riego, M.V., Fernandez, R.M., Roman-Rodriguez, F.J., Moya-Jimenez, M.J., de Agustin, J.C., Antinolo, G., Borrego, S., 2014, "*Involvement of DNMT3B in the pathogenesis of Hirschsprung disease and its possible role as a regulator of neurogenesis in the human enteric nervous system*", Genetics in medicine : official journal of the American College of Medical Genetics 16, 703-710.

Touré, A.M., Landry, M., Souchkova, O., Kembel, S.W., Pilon, N., 2019, "*Gut microbiota-mediated Gene-Environment interaction in the TashT mouse model of Hirschsprung disease*", Scientific reports 9, 492.

Troffer-Charlier, N., Doerflinger, N., Metzger, E., Fouquet, F., Mandel, J.L., Aubourg, P., 1998, "*Mirror expression of adrenoleukodystrophy and adrenoleukodystrophy related genes in mouse tissues and human cell lines*", European journal of cell biology 75, 254-264.

Uddin, B., Chen, N.P., Panic, M., Schiebel, E., 2015, "*Genome editing through large insertion leads to the skipping of targeted exon*", BMC genomics 16, 1082.

Uesaka, T., Nagashimada, M., Enomoto, H., 2013, "*GDNF signaling levels control migration and neuronal differentiation of enteric ganglion precursors*", The Journal of neuroscience : the official journal of the Society for Neuroscience 33, 16372-16382.

Underwood, C.K., Coulson, E.J., 2008, "*The p75 neurotrophin receptor*", Int J Biochem Cell Biol 40, 1664-1668.

Usui, N., Co, M., Harper, M., Rieger, M.A., Dougherty, J.D., Konopka, G., 2017, "*Sumoylation of FOXP2 Regulates Motor Function and Vocal Communication Through Purkinje Cell Development*", Biological psychiatry 81, 220-230.

van den Berghe, P.V., **Klomp**, L.W., 2009, "*New developments in the regulation of intestinal copper absorption*", Nutr Rev 67, 658-672.

Wallace, A.S., **Anderson**, R.B., 2011, "*Genetic interactions and modifier genes in Hirschsprung's disease*", World journal of gastroenterology 17, 4937-4944.

Wallace, A.S., **Burns**, A.J., 2005, "*Development of the enteric nervous system, smooth muscle and interstitial cells of Cajal in the human gastrointestinal tract*", Cell and tissue research 319, 367-382.

Wang, F., Flanagan, J., Su, N., Wang, L.C., Bui, S., Nielson, A., Wu, X., Vo, H.T., Ma, X.J., Luo, Y., 2012, "*RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues*", The Journal of molecular diagnostics : JMD 14, 22-29.

Wang, W., Chen, Y., Wang, S., Hu, N., Cao, Z., Wang, W., Tong, T., Zhang, X., 2014, "*PIASxalpha ligase enhances SUMO1 modification of PTEN protein as a SUMO E3 ligase*", The Journal of biological chemistry 289, 3217-3230.

Watanabe, M., Itoh, K., 2011, "*Characterization of a novel posttranslational modification in neuronal nitric oxide synthase by small ubiquitin-related modifier-1*", Biochimica et biophysica acta 1814, 900-907.

Westphal, D., Dewson, G., Czabotar, P.E., Kluck, R.M., 2011, "*Molecular biology of Bax and Bak activation and action*", Biochimica et biophysica acta 1813, 521-531.

Wilentz, R.E., Witters, L.A., Pizer, E.S., 2000, "*Lipogenic enzymes fatty acid synthase and acetyl-coenzyme A carboxylase are coexpressed with sterol regulatory element binding protein and Ki-67 in fetal tissues*", Pediatr Dev Pathol 3, 525-531.

Wu, W., Lu, L., Xu, W., Liu, J., Sun, J., Zheng, L., Sheng, Q., Lv, Z., 2018, "*Whole Exome Sequencing Identifies a Novel Pathogenic RET Variant in Hirschsprung Disease*", Frontiers in genetics 9, 752.

Yang, Y., He, Y., Wang, X., Liang, Z., He, G., Zhang, P., Zhu, H., Xu, N., Liang, S., 2017, "*Protein SUMOylation modification and its associations with disease*", Open Biol 7.

Young, H.M., Hearn, C.J., Farlie, P.G., Canty, A.J., Thomas, P.Q., Newgreen, D.F., 2001, "*GDNF is a chemoattractant for enteric neural cells*", Developmental biology 229, 503-516.

Yunus, A.A., **Lima**, C.D., 2009, "*Structure of the Siz/PIAS SUMO E3 ligase Siz1 and determinants required for SUMO modification of PCNA*", Molecular cell 35, 669-682.

Zhang, G., 2019, "*Insights into mitotic checkpoint by integrating CRISPR and RNAi*", Mol Cell Oncol 6, 1603436.

Zhang, X.H., Tee, L.Y., Wang, X.G., Huang, Q.S., Yang, S.H., 2015, "*Off-target Effects in CRISPR/Cas9-mediated Genome Engineering*", Molecular therapy. Nucleic acids 4, e264.

Zhang, Z., Li, Q., Diao, M., Liu, N., Cheng, W., Xiao, P., Zou, J., Su, L., Yu, K., Wu, J., Li, L., Jiang, Q., 2017, "*Sporadic Hirschsprung Disease: Mutational Spectrum and Novel Candidate Genes Revealed by Next-generation Sequencing*", Scientific reports 7, 14796.

Zhou, W., Ye, S.D., Chen, C., Wang, W., 2018, "*Involvement of RBP4 in Diabetic Atherosclerosis and the Role of Vitamin D Intervention*", Journal of diabetes research 2018, 7329861.

Ziegler, A.B., Thiele, C., Tenedini, F., Richard, M., Leyendecker, P., Hoermann, A., Soba, P., Tavosanis, G., 2017, "*Cell-Autonomous Control of Neuronal Dendrite Expansion via the Fatty Acid Synthesis Regulator SREBP*", Cell reports 21, 3346-3353.

Unpublished work

Bachelor Thesis **Dawid**, D., 2019; "*Identification and characterization of novel candidate genes in familial Hirschsprung´s disease* ", Faculty of Biosciences/Faculty of Chemistry and Earth Sciences; Ruprecht-Karls University Heidelberg; Bachelor of Science;

Master Thesis **Mederer**, T., 2015; "*Molecular characterisation of enteric precursor cells during development*", Faculty of Biosciences; Ruprecht-Kalrs University Heidelberg; Master of Science;

Master Thesis **Volz**, J., 2014; "*Identification and characterization of genes involved in Hirschsprung's disease - a developmental disorder of the enteric nervous system*", Faculty of Biosciences; Ruprecht-Karls-University Heidelberg; Master of Science;

Turner, T., Nguyen, K. D., Krumm, N., Chatterjee, S., Kapoor, A., Jiang, Q., Ling, A. Y., Sosa, M. X., Gupta, N., Eichler, E. E., Gabriel, S., Berrios, C., Chakravarti, A., 2013, "*The burden of coding, non-coding and chromosomal mutations in Hirschsprung disease.*" , Abstract, Presented at the Annueal meeting of the American Society of Human Genetics (ASHG), Boston.

Webpages

Howard Hughes Medical Institute; 30.06.19; https://www.hhmi.org/content/zhang-yi-research-abstract-slideshow

gnomAD; 30.06.19; https://gnomad.broadinstitute.org/about

NCBI; 30.06.19; *Gene ID: 215, ABCD1*; https://www.ncbi.nlm.nih.gov/gene/215

NCBI; 30.06.19; *Gene ID: 538, ATP7A*; https://www.ncbi.nlm.nih.gov/gene/?term=538

NCBI; 30.06.19; *Gene ID: 1889, ECE1*; https://www.ncbi.nlm.nih.gov/gene/?term=1889

NCBI; 30.06.19; *Gene ID: 2055, CLN8*; https://www.ncbi.nlm.nih.gov/gene/?term=2055

NCBI; 30.06.19; *Gene ID: 5274, SERPINI1*; https://www.ncbi.nlm.nih.gov/gene/?term=5274

NCBI; 30.06.19; *Gene ID: 5950, RBP4*; https://www.ncbi.nlm.nih.gov/gene/5950

NCBI; 30.06.19; *Gene ID: 6720, SREBF1*; https://www.ncbi.nlm.nih.gov/gene/6720

NCBI; 30.06.19; *Gene ID: 9063, PIAS2*; https://www.ncbi.nlm.nih.gov/gene/9063

NCBI; 30.06.19; *Gene ID: 51684, SUFU*; https://www.ncbi.nlm.nih.gov/gene/?term=51684

NCBI; 30.06.19; *Gene ID: 56115, PCDHG*; https://www.ncbi.nlm.nih.gov/gene/56115

NCBI; 30.06.19; *Gene ID: 89797, NAV2*; https://www.ncbi.nlm.nih.gov/gene/89797

NCBI; 30.06.19; *Gene ID: 100507436, MICA*; https://www.ncbi.nlm.nih.gov/gene/100507436

OMIM; 30.06.19; # 600143; *CLN8*; https://www.omim.org/entry/607837

The Human Protein Atlas; 30.06.19; *ABCD1*; https://www.proteinatlas.org/ENSG00000101986-ABCD1/tissue

The Human Protein Atlas; 30.06.19; *MICA*; https://www.proteinatlas.org/ENSG00000204520-MICA/tissue

The Human Protein Atlas; 30.06.19; *UHRF1BP1L*; https://www.proteinatlas.org/ENSG00000111647-UHRF1BP1L/tissue

The Human Protein Atlas; 30.06.19; *CLN8*; https://www.proteinatlas.org/ENSG00000182372-CLN8/tissue

ACKNOWLEDGEMENTS

First, I would like to thank **Prof. Dr. Beate Niesler** for the opportunity to pursue my doctorate in her lab. Thank you for being a great supervisor, supporting me at any moment and always having an open ear for any kind of problems or challenges. I highly appreciated your positive thinking, patience, humour and enthusiasm throughout this time and I learned a lot from your optimism.

I further would like to thank **Prof. Dr. Gudrun Rappold** for being my supervisor and her continuous support in various ways and on different levels throughout my PhD. I highly appreciated your help.

I would like to thank my TAC members **Prof. Dr. Hilmar Bading** and **PD. Dr. Marco Metzger** for critical discussions of the project and their constructive and very helpful suggestions in the context of my TAC meetings. Further, I would like to thank **Prof. Dr. Hilmar Bading** and **Dr. Moritz Mall** for accepting of being my examiners.

I am very grateful for direct supervision of **Dr. Stefanie Schmitteckert** who never refused trying to uncover one of my countless mysteries – at Friday afternoon or at any other time of the week. You were a great support with all your experience, your creativity and your optimism and I highly appreciated to work with you. You definitely deserve a lab octopus (with at least eight brains and arms)!

Special thanks as well to all people that contributed at various sites to the project:
First of all, to **Julia Volz** and all people involved in the initiation of the project as these analyses provided the basis for my PhD thesis.
To all **patients** and all people from the Pediatrics surgery Heidelberg, especially, **Dr. Philipp Romero** and **Leonie Carstensen** who were always very motivated and supportive. Thanks to their efforts, the project can further grow and develop.
To **Dr. Nagarajan Paramasivam** who was always approachable for support in all bioinformatic related issues and tried his best to explain it in a ‘non-bioinformaticians’ way.
To **Dr. Cornelia Thöni** and **Jutta Scheuerer** for their quick and straightforward help with regard to the human archived specimens and respective analyses.
To **Dr. Thomas Thumberger** for introducing us to the ‘CRISPR-world’ and his continuous willingness to explain or to support when challenges arose.
To **Dr. Melanie Bewerunge-Hudler** for her support in the microarray analyses and to **Dr**. **Carolina de la Torre** who finally enabled us to successfully conclude our little “microarray-odyssey”.
To **Dr. Volker Eckstein** and the team of the FACS core facility for FACS analyses whenever needed.
To all **members of the International Hirschsprung Consortium** who shared the data of their HSCR patients with us.
To **Prof. Dr. Axel Schweickert** and **Tim Felger** who provided the opportunity to follow up our “zombie” protein in another experimental setup as all our approaches failed.
To **Dr. Carlo Beretta** from the Cell Networks Math-Clinic who showed me how to make image analysis at least a little bit easier when using the right tools in ImageJ.

To **Prof. Dr. Justo Lorenzo Bermejo** from the Institute of Medical Biometry and Informatics for his quick and straightforward advice on how to do statistical data analysis without making it too complicated.

Thanks to **Studienstiftung des deutschen Volkes e.V.** - as I would not have been able to work on this PhD project without respective funding- and to the **HBGIS**. I profited in various ways from the respective support.

Furthermore, I would like to thank all former and current members of the **AG Niesler** and **AG Rappold**. You were a great team and created a very pleasant working atmosphere which I highly appreciated.
Special thanks to **Ralph Röth**, **Birgit Weiß** and **Inge Bender** for their willingness to share their own experiences, giving advice whenever needed, supporting in various ways and like this making everyone's life much easier.
Furthermore, I would like to thank **Heike Kuzan** and **Verena Wahl** for their general help in the organization of the day-to-day work in lab 521 and more specifically for their support in my project whenever needed.
Thanks to **Antonino Montalbano** and **Ana de Sena Cortabitarte** who introduced me into the 'lab world' of the 5th floor and helped whenever I was lost in protocols or kits or both at the same time.
Further, I would like to thank the MD students **Valerie Linz** and **Christine Unsicker** – not only for cheering me up with all their questions but as well for becoming really good friends. I highly appreciated to work with you.
Moreover, thanks to **Denise Dawid** for her help in the final phase of my project - and for providing a lot of sweets for the "Diabetes drawer".

I would like to dedicate a special thank you to **Flavia-Bianca Cristian** and **Simon Sumer**. You made a huge difference to me in the lab throughout all the time – not only because of your scientific advice and practical support whenever needed but as well because you saved me from several panic attacks and nervous breakdowns. We shared so much chocolate, cookies & Co. and had so many depressed and funny moments, which I will never forget…I learned a lot from the both of you!

Further, I would like to thank **Linda Manhart** and **Colin Lischik** – we have started all of this together in Mainz and continued in Heidelberg… at some point, with some spatial separation. During the last four years, we met for several lunch breaks, where one of us (mostly Linda or me…) usually was even more depressed than the two others. However, we always found something/someone to laugh about. I highly appreciated your input (not only on scientific level). And finally, we all made it…!!!

And last, I would like to thank my parents, my sister and my friends. Thank you for everything you have done for me: for accepting me the way I am, for your patience and your encouragement during the last years. I am deeply grateful that I can count on you any time.

DECLARATION OF ACADEMIC INTEGRITY

SWORN AFFIDAVIT ACCORDING TO §8 OF THE DOCTORAL DEGREE REGULATIONS OF THE COMBINED FACULTY OF NATURAL SCIENCES AND MATHEMATICS

I herewith declare that:

1. The thesis I have submitted entitled “Identification and characterization of genes involved in Hirschsprung’s disease- a developmental disorder of the enteric nervous system” is my own work. I have only used the sources indicated and have not made unauthorized use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given.

2. I have not yet presented this thesis or parts thereof to a university as part of an examination or degree.

3. I am aware of the importance of a sworn affidavit and the criminal prosecution in case of a false or incomplete affidavit.

I affirm that the above is the absolute truth to the best of my knowledge and that I have not concealed anything.

Place and date

Signature