



Lab Resource: Multiple Cell Lines

Generation of two induced pluripotent stem cell lines from a female adult homozygous for the Wilson disease associated ATP7B variant p.H1069Q (AKOSi008-A) and a healthy control (AKOSi009-A)

Janine Petters^a, Christin Völkner^a, Saskia Krohn^b, Hugo Murua Escobar^b, Jörn Bullerdiek^c, Ulrike Reuner^d, Moritz J. Frech^{a,e}, Andreas Hermann^{a,e,f}, Jan Lukas^{a,e,*}

^a Translational Neurodegeneration Section Albrecht-Kossel, Department of Neurology, University Medical Center Rostock, 18147 Rostock, Germany

^b Department of Medicine, Clinic III – Hematology, Oncology, Palliative Medicine, University Medical Center Rostock, 18057 Rostock, Germany

^c Institute of Medical Genetics, University Medical Center Rostock, University of Rostock, 18057 Rostock, Germany

^d Department of Neurology, Technische Universität Dresden, 01307 Dresden, Germany

^e Center for Transdisciplinary Neurosciences Rostock (CTNR), University Medical Center Rostock, University of Rostock, 18147 Rostock, Germany

^f German Center for Neurodegenerative Diseases (DZNE) Rostock/Greifswald, 18147 Rostock, Germany

A B S T R A C T

Wilson disease (WD) is a rare, monogenic disorder caused by mutations in the gene *ATP7B*. A loss of function of the expressed protein leads to excessive hepatic and cerebral copper storage. In this study, we present the generation of two induced pluripotent stem cell (iPSC) lines derived from fibroblasts of a clinically asymptomatic, chelator treated female WD patient carrying the common missense mutation p.H1069Q and an age-matched female healthy control subject. The generated iPSC lines expressed pluripotency markers, showed differentiation potential and retained their parental genotype. Therefore, these cells provide a valuable resource to understand the pathophysiology of WD and can be used as model systems for drug testing.

Resource Table

Unique stem cell lines identifier	1: AKOSi008-A 2: AKOSi009-A
Alternative names of stem cell lines	1: iPS D1-12 2: iPS GM23251-4
Institution	Translational Neurodegeneration Section “Albrecht-Kossel”, Department of Neurology, University Medical Center Rostock, University of Rostock, 18147 Rostock, Germany
Contact information of distributor	Dr. Jan Lukas; jan.lukas@med.uni-rostock.de
Type of cell lines	iPSC
Origin	Human
Cell Source	Fibroblasts
Clonality	Clonal
Method of reprogramming	Retrovirus
Multiline rationale	Age- and sex-matched Wilson disease and healthy control cell lines
Gene modification	Yes
Type of modification	Hereditary
Associated disease	Wilson disease
Gene/locus	1: <i>ATP7B</i> / 13q.14.3c.3207C>A/c.3207C>A
Method of modification	N/A

(continued on next column)

Resource Table (continued)

Name of transgene or resistance	OCT4, SOX2, KLF4, C-MYC
Inducible/constitutive system	N/A
Date archived/stock date	1: July 2019 2: April 2019
Cell line repository/bank	N/A
Ethical approval	Fibroblasts GM23251 were obtained from the NIGMS Human Genetic Cell Repository at the Coriell Institute for Medical Research, USA. Fibroblasts D1-12 were obtained in house by skin biopsy derived from a female Wilson disease patient. All procedures were in accordance with the Helsinki convention and approved by the Ethical Committee of the University of Dresden (EK45022009; EK393122012).

1. Resource utility

Mutations in the gene *ATP7B* lead to the rare disorder of copper metabolism Wilson disease (WD). Due to the prominent role of the p.

* Corresponding author.

<https://doi.org/10.1016/j.scr.2020.102079>

Received 16 October 2020; Accepted 1 November 2020

Available online 5 November 2020

1873-5061/© 2020 The Authors.

Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

H1069Q variant in the European WD patient cohort, human iPSC lines carrying this mutation are suitable for disease phenotyping as well as for individualized therapy development for a significant proportion of WD patients.

2. Resource

Patient-specific induced pluripotent stem cells (iPSCs) display the patient's individual genetic background and, when differentiated into a disease-relevant cell type, are suitable model systems for genetic diseases. The development of such cell models is crucial for rare diseases for which a clear genotype/phenotype correlation does not yet exist, such as Wilson disease (WD), an autosomal recessive disorder of copper metabolism (Ferenzi et al., 2019). For this reason, the generation of human, homozygous disease models is of particular importance for the representation of pathophysiology and for the development of therapeutic drugs that depend on the patients' genetics. Here, we generated a WD-specific and a healthy control iPSC line. Both cell lines were sex- and age-matched to obtain a high comparability. Fibroblasts from two female individuals were reprogrammed into iPSC lines using retroviral vectors for OCT4, KLF4, SOX2 and C-MYC. Cell line D1 (Fibro 008-A) represented fibroblasts from a WD patient homozygous for the p. H1069Q (c.3207C>A) variant. This patient was diagnosed during family diagnostics at asymptomatic stage and since then treated using classical chelation therapy with D-Penicillamine (1200 mg/day), vitamin B6 and vitamin D. At the time of skin biopsy she was clinically still unaffected. GM23251 (Fibro 009-A, Coriell Institute for Medical Research) was a healthy control fibroblast line. Successful transduction of the parental fibroblasts led to the formation of iPSC colonies. Before colonies were picked and expanded on irradiated mouse embryonic fibroblasts (MEF), retroviral silencing was determined by the loss of GFP signal in the iPSCs (Supplementary Fig. S1A). In the further course of passaging, the silencing of the virally expressed *OCT4*, *SOX2*, *KLF4* and *C-MYC* transcription factors was verified by specific reverse transcription PCR (RT-PCR) (Supplementary Fig. S1B) for iPS 008-A at passage 11 and iPS 009-A at passage 20. One clone of each line was selected for pluripotency characterization (Tables 1 and 2). Clones iPS AKOSi008-A (iPS 008-A) and iPS AKOSi009-A (iPS 009-A) showed a normal stem cell-like morphology (Fig. 1A). The activity of alkaline phosphatase could be detected in the iPSCs, but not in the MEF feeder cells. The colorimetric test resulted in a dark purple staining of alkaline phosphatase positive cells (Fig. 1B). For further characterization, both iPSC lines were passaged on Matrigel-coated plates. Immunofluorescent staining proved expression of pluripotency-associated surface and intracellular markers, such as OCT4, NANOG, TRA-1-60 and TRA-1-81 (Fig. 1C). Quantitative flow cytometry analysis revealed high levels of expressed pluripotency markers in both iPS 008-A and iPS 009-A (Fig. 1D and Table 2). Additionally, selected pluripotency genes were tested using RT-PCR. The presence of significant levels of mRNAs for *OCT4*, *NANOG*, *KLF4*, *SOX2*, *C-MYC*, *hTERT*, *ZFP296*, *FGF4* and *ESG1* could be observed in the iPSC lines (Fig. 1E). Little to no signal was detected in the parental fibroblasts. *GAPDH* was used as control. *OCT4*, *SOX2*, *KLF4* and *C-MYC* primer pairs specifically amplified endogenous transcripts (Table 3). The formation of embryoid bodies demonstrated the ability of the iPSC lines to differentiate into cells of all three germ layers. Immunocytochemistry proved positive staining for alpha-fetoprotein (aFP, endoderm), nestin (ectoderm) and muscle actin (MA, mesoderm) in iPS 008-A and iPS 009-A (Fig. 1F). Chromosome analysis showed normal karyotypes (46, XX) for iPS 008-A at passage 12 and iPS 009-A at passage 18 (Fig. 1G). Both

Table 2
Characterization and validation.

Classification	Test	Result	Data
Morphology	Photography	Visual record of the line: Normal	Fig. 1 panel A
Phenotype	Qualitative analysis: Alkaline phosphatase staining	Positive	Fig. 1 panel B
	Qualitative analysis: Immunocytochemistry	Expression of pluripotency markers: OCT4, NANOG, SSEA4, TRA-1-60, TRA-1-81	Fig. 1 panel C (representative)
	Qualitative analysis: RT-PCR	Expression of pluripotency genes: <i>OCT4</i> , <i>NANOG</i> , <i>KLF4</i> , <i>SOX2</i> , <i>C-MYC</i> , <i>hTERT</i> , <i>ZFP296</i> , <i>FGF4</i> and <i>ESG1</i>	Fig. 1 panel E
	Quantitative analysis: Flow cytometry	Percentage of positive cells iPS 008-A: OCT4: 96.0%, NANOG: 98.0%, SSEA4: 96.1%, TRA-1-60: 91.7%, TRA-1-81: 88.11% iPS 009-A: OCT4: 97.6%, NANOG: 98.7%, SSEA4: 97.2%, TRA-1-60: 95.8%, TRA-1-81: 93.4%	Fig. 1 panel D (representative)
Genotype	Karyotype (G-banding) and resolution	46, XX Resolution: 300–550	Fig. 1 panel G
Identity	Microsatellite PCR (mPCR) OR STR analysis	Not performed 18 STR loci tested, 100% matched	N/A Available with authors
Mutation analysis (IF APPLICABLE)	Sequencing	iPS 008-A: homozygous p. H1069Q	Supplementary Fig. S1 panel C
Microbiology and virology	Southern Blot OR WGS Mycoplasma	Not performed Mycoplasma testing by PCR: Negative	N/A Supplementary Fig. S1 panel D
Differentiation potential	Embryoid body formation	Expression of genes in embryoid bodies: Muscle actin (MA), nestin and α -fetoprotein (aFP)	Fig. 1 panel F
Donor screening (OPTIONAL)	HIV 1 + 2 Hepatitis B, Hepatitis C	Not performed	N/A
Genotype additional info (OPTIONAL)	Blood group genotyping	Not performed	N/A
	HLA tissue typing	Not performed	N/A

fibroblast and iPSC lines were sequenced using targeted NGS sequencing. The homozygous p.H1069Q mutation could be detected in iPS 008-A as well as in the corresponding fibroblasts Fibro 008-A

Table 1
Summary of lines.

iPSC line names	Abbreviation in figures	Gender	Age	Ethnicity	Genotype of locus	Disease
AKOSi008-A (iPS D1-12)	iPS 008-A	Female	54	Caucasian	NM_000053.4:c.3207C>A/hom	Wilson disease
AKOSi009-A (iPS GM23251-4)	iPS 009-A	Female	50	Caucasian	N/A	None

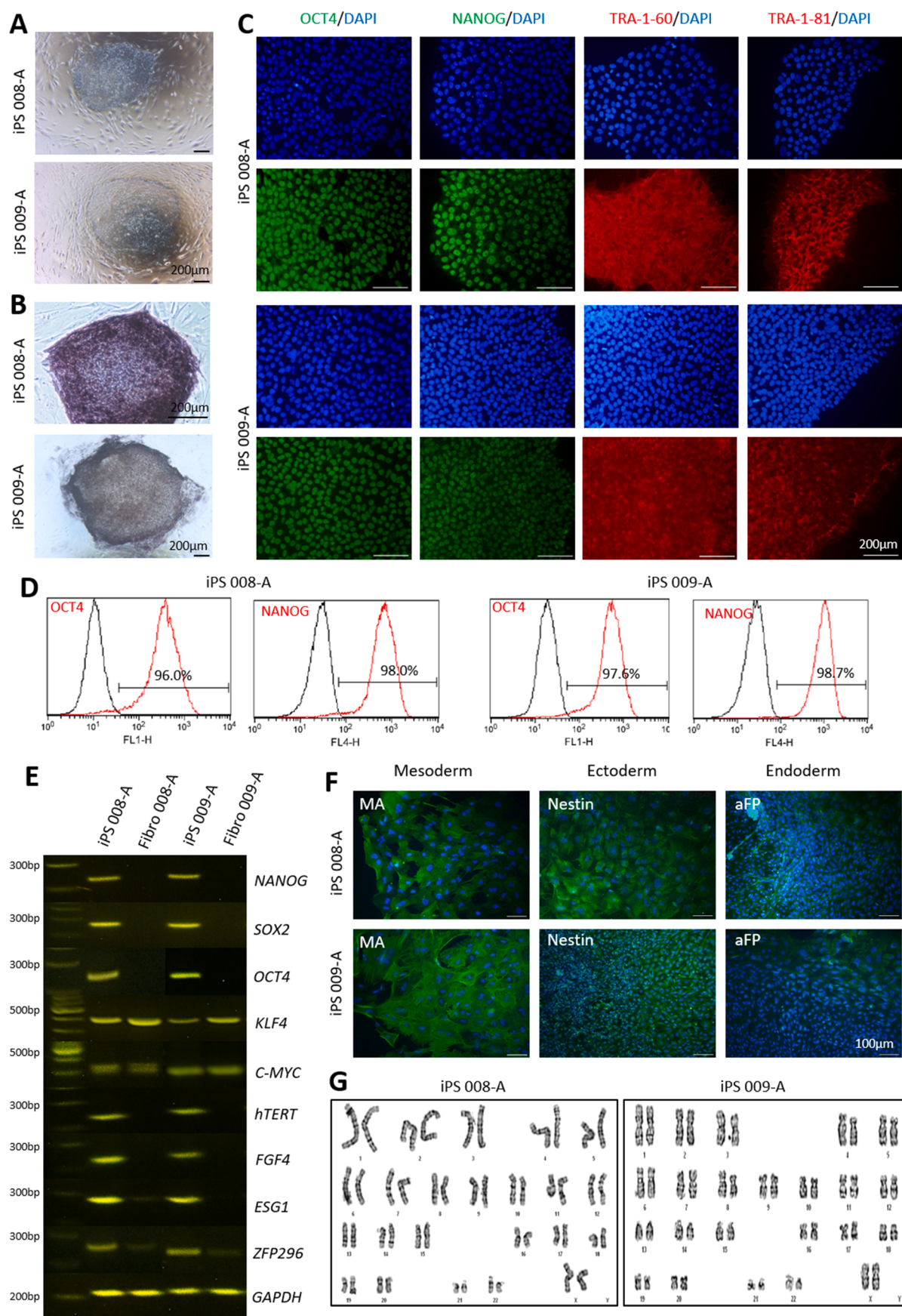


Fig. 1. Characterization of induced pluripotent stem cell lines AKOSi008-A and AKOSi009-A from fibroblasts of a WD patient and a healthy control, respectively.

Table 3
Reagents details.

Antibodies used for immunocytochemistry/flow-cytometry			
	Antibody	Dilution	Company Cat # and RRID
Pluripotency Marker (IF)	Rabbit anti-OCT4	1:100	Stemgent Cat# 09-0023, RRID: AB_2167689
Pluripotency Marker (IF)	Rabbit anti-NANOG	1:100	Stemgent Cat# 09-0020, RRID: AB_2298294
Pluripotency Marker (IF)	Mouse anti-SSEA4	1:100	Stemgent Cat# 09-0006, RRID: AB_1512169
Pluripotency Marker (IF)	Mouse anti-TRA-1-60	1:100	Stemgent Cat# 09-0010, RRID: AB_1512170
Pluripotency Marker (IF)	Mouse anti-TRA-1-81	1:100	Stemgent Cat# 09-0011, RRID: AB_1512171
Pluripotency Marker (FC)	Alexa Fluor 488 anti-OCT4, mouse IgG2b	1:20	BioLegend Cat# 653705, RRID: AB_2562250
Pluripotency Marker (FC)	Alexa Fluor 647 anti-NANOG, mouse IgG1	1:50	BioLegend Cat# 674210, RRID: AB_2650619
Pluripotency Marker (FC)	Alexa Fluor 647 anti-SSEA-4, mouse IgG3	1:500	BioLegend Cat# 330407, RRID: AB_1089201
Pluripotency Marker (FC)	PE anti-human TRA-1-60-R, mouse IgM	1:20	BioLegend Cat# 330609, RRID: AB_1279447
Pluripotency Marker (FC)	Alexa Fluor 488 anti-TRA-1-81, mouse IgM	1:20	BioLegend Cat# 330709, RRID: AB_2561741
Differentiation Marker (IF)	Mouse anti-Muscle actin	1:50	Agilent Dako Cat# M0635, RRID: AB_2242301
Differentiation Marker (IF)	Mouse anti-Nestin	1:100	R and D Systems Cat# MAB1259, RRID: AB_2251304
Differentiation Marker (IF)	Mouse anti-Alpha fetoprotein	1:20	R and D Systems Cat# MAB1368, RRID: AB_357658
Secondary antibody	Alexa Fluor 488, Goat anti-mouse IgG	1:500	Thermo Fisher Scientific Cat# A-11029, RRID: AB_2534088
Secondary antibody	Alexa Fluor 568, Goat anti-mouse IgM	1:500	Thermo Fisher Scientific Cat# A-21043, RRID: AB_2535712
Secondary antibody	Alexa Fluor 488, Goat anti-rabbit IgG	1:500	Thermo Fisher Scientific Cat# A-11034, RRID: AB_2576217
Primers			
	Target	Forward/Reverse primer (5'-3')	
Pluripotency marker (RT-PCR)	<i>C-MYC</i>	GCGTCCTGGGAAGGGAGATCCGAGC/ TTGAGGGGCATCGTCGGGGAGGCTG	
Pluripotency marker (RT-PCR)	<i>NANOG</i>	TGTTGTTCTCTCCACCCAGC/ACCAGGTCTTCACCTGTTTGT	
Pluripotency marker (RT-PCR)	<i>OCT4</i>	GACAGGGGGAGGGAGGAGCTAGG/ CTTCCTCCCAACCAAGTTGCCCAAAAC	
Pluripotency marker (RT-PCR)	<i>SOX2</i>	AGGGAGAGAAGTTTGAGCCC/GCAGAGAAAATCAGGCGAAG	
Pluripotency marker (RT-PCR)	<i>KLF4</i>	ACGATCGTGGCCCGGAAAAGGACC/ TGATTGTAGTGCTTTCTGGCTGGGCTCC	
Pluripotency marker (RT-PCR)	<i>ZFP296</i>	CTGGACCGACAAACACCCAG/CTTCAGCTCCTCTCGTTCTGAG	
Pluripotency marker (RT-PCR)	<i>ESG1</i>	ATATCCCGCCGTGGGTGAAAGTTC/ ACTCAGCCATGGACTGGAGCATCC	
Pluripotency marker (RT-PCR)	<i>FGF4</i>	CAAGCTCTATGGCTCGCCCT/TCCTCCCATCTTGTCTCAGGG	
Pluripotency marker (RT-PCR)	<i>hTERT</i>	GAGCTGACGTGGAAGATGAGC/CATCAGCCAGTCGACGGAACCT	
House-Keeping Gene (RT-PCR)	<i>GAPDH</i>	CATGTTCCAATATGATTCCACCC/GGGATCTCGCTCCTGGAAGAT	
Viral OCT4 expression (RT-PCR)	<i>OCT4_pMIG/IRES</i>	GTACTCCTCGGTCCCTTTCC/GCATTCTTTGGCGGAGAG	
Viral SOX2 expression (RT-PCR)	<i>SOX2_pMIG/IRES</i>	CATGTCCAGCACTACCAGA/GCATTCTTTGGCGGAGAG	
Viral KLF4 expression (RT-PCR)	<i>KLF4_pMIG/IRES</i>	CCCACACAGGTGAGAAACCT/GCATTCTTTGGCGGAGAG	
Viral C-MYC expression (RT-PCR)	<i>C-MYC_MSCV/IRES</i>	AAGAGGACTTGTTCGGGAAA/GCATTCTTTGGCGGAGAG	

(Supplementary Fig. S1C). In cell control line iPS 009-A and their corresponding fibroblasts Fibro 009-A no pathogenic *ATP7B* variant was identified. Short tandem repeat (STR) analysis of 18 genomic loci was performed and revealed 100% identity between fibroblasts and the respective iPS line (archived at journal). As a quality standard, mycoplasma testing was performed. No contamination could be detected in the iPS lines (Supplementary Fig. S1D).

3. Materials and methods

3.1. Derivation of fibroblasts

For generation of iPS lines, fibroblast lines were established from skin biopsies taken from the Wilson disease patient as described previously (Naumann et al., 2018).

3.2. Reprogramming of fibroblasts

Retroviral reprogramming of fibroblasts using vectors for OCT4, SOX2, KLF4 and C-MYC was performed as described elsewhere (Peter et al., 2017).

3.3. Cell culture

iPS cells on MEF feeder cells were cultured in iPS medium (DMEM/F12, 20% Knockout serum replacement, 0.1 mM MEM non-essential amino acids, 1% GlutaMax, 0.1 mM 2-mercaptoethanol, 1% penicillin-streptomycin (all Gibco) supplemented with 12.5 ng/ml FGF-2 (Amsbio)) and passaged mechanically for the first 4–10 passages. After iPS cells were transferred onto Matrigel-coated plates (Corning), the cells were maintained in mTeSR1 medium (STEMCELL Technologies) with 0.25% penicillin-streptomycin (10000 U/ml, Gibco). iPS cells on Matrigel were

passed at a split ratio of 1:4–1:6 using Dispase (STEMCELL Technologies) for routine passaging and Gentle Cell Dissociation Reagent (STEMCELL Technologies) for single cell suspensions. Fibroblasts and MEF feeder cells were cultured in high glucose DMEM (Gibco) supplemented with 10% fetal bovine serum (FBS, GE Healthcare) and 1% penicillin-streptomycin (Gibco). All cells were incubated in a humidified incubator at 37 °C and 5% CO₂.

3.4. Alkaline phosphatase (AP) activity test

On the day of splitting, iPS colonies on MEF feeder cells were fixed in precooled methanol for 10 min and incubated with AP staining solution (75% distilled water, 10% 1 M sodium chloride, 10% 1 M Tris (pH 9.8), 5% 1 M magnesium chloride, 1:50 NBT/BCIP stock solution (Roche)) for 15 min. Color change was detected using a Nikon Eclipse TS100 microscope.

3.5. Karyotyping

Cells were incubated with fresh mTeSR1 medium for 2 h and treated with 0.15 µg/ml colcemid (KaryoMAX, Gibco) for 90 min at 37 °C. After treatment, cells were harvested with Accutase (Sigma-Aldrich). The cell pellet was resuspended in 0.075 M potassium chloride and incubated for another 30 min at 37 °C and subsequently fixed in methanol-acetic acid solution (3:1). Chromosome analysis was performed using G-banding technique. At least 18 metaphases were counted for each iPS line.

3.6. Immunocytochemistry

iPS colonies were seeded onto Matrigel-coated glass cover slips. After 3–5 days, cells were fixed in 4% paraformaldehyde for 15 min at room temperature. Blocking (10% normal goat serum (NGS), 0.1%

Triton-X 100 in PBS) was performed for 45 min at room temperature followed by primary antibody incubation overnight at 4 °C and secondary antibody incubation for 1 h at room temperature on the next day. Antibodies and antibody dilutions are listed in Table 3. DAPI (250 ng/ml) was added for 5 min at room temperature. After fixation, blocking, antibody and DAPI incubations, cells were washed twice with PBS. A digital compact microscope (Keyence) was used for imaging.

3.7. RT-PCR

Total RNA was isolated using Quick-RNA Miniprep kit (Zymo Research) following the manufacturer's protocol. RNA concentration was measured in a multifunction reader equipped with NanoQuant plate (Tecan). One-step Reverse Transcriptase PCR (QIAGEN) was performed in an Eppendorf 5331 MasterCycler Gradient Thermal Cycler. 50 ng of total RNA were used for each reaction. PCR products were run on 1.5% – 2.5% TBE agarose gels. Bands were detected with an UV transilluminator (Herolab). For the verification of viral vector silencing, the PCR was performed for 30 cycles using the respective viral vector as positive control. Used primers are listed in Table 3.

3.8. Flow cytometry

Fixation and permeabilization for intracellular stainings were done using True-NuclearTM Transcription Factor Buffer Set (BioLegend). Cells for surface marker analysis were blocked with PBS + 10% FBS (GE Healthcare) for 30 min on ice. For both surface and intracellular stainings, cells were incubated with fluorophore-conjugated antibodies for 90 min at room temperature. Antibodies and antibody dilutions are listed in Table 3. FACSCalibur (BD) was used for measurement. 5x10⁴ cells per sample were analyzed. Data analysis was performed with software FCSalyzer version 0.9.18-alpha (<https://sourceforge.net/projects/fcsalyzer/>).

3.9. Embryoid body (EB) formation

IPSC colonies were mechanically detached from the MEF feeder cell layer, transferred to a low attachment plate and incubated at 37 °C and 5% CO₂ for 5 days. IPSC medium was changed to EB medium (78% Knockout DMEM, 0.1 mM MEM non-essential amino acids, 1% Gluta-Max, 0.1 mM 2-mercaptoethanol, 0.25% penicillin–streptomycin (all Gibco) and 20% FBS (GE Healthcare)). Following this, EBs were seeded onto gelatin-coated cover slips and cultivated for 9 days. EBs were fixed in 4% paraformaldehyde, permeabilized in 0.1% Triton-X 100 and separately blocked with 4% NGS. Antibodies that were used for spontaneous differentiation are listed in Table 3.

3.10. Short tandem repeat (STR) analysis

Cell suspensions of fibroblasts and iPSCs at a concentration of 1x10⁶ cells/ml were collected on Whatman® FTA® cards according to the manufacturer's protocol (ATCC, VA, USA). 18 STR loci were analyzed and compared by ATCC.

3.11. Targeted sequencing

Genomic DNA was extracted (Quick-DNATM Miniprep Kit, Zymo

Research) and quantified using a multifunction reader (Tecan). Targeted sequencing library construction was performed using a custom designed Ion AmpliSeqTM ATP7B Panel (Thermo Fisher Scientific). 10 ng of genomic DNA were used to amplify the complete coding sequence. Sequencing was carried out on an Ion TorrentTM Personal Genome MachineTM System, using an Ion Torrent 318 V2 chip. Sequence check was performed against the hg19 assembly of the human genome using Torrent SuiteTM software and the variant caller plugin version 5.10.0.18 (Thermo Fisher Scientific).

3.12. Mycoplasma detection

Mycoplasma detection was carried out using PCR Mycoplasma Test Kit I/C (PromoCell) following manufacturer's instructions.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank Matthew L. Chang for technical support and Heike Janeczek-Meyer (Department of Medical Genetics, University Medical Center Rostock) for technical assistance with chromosome analysis.

Funding information

This work was supported by the European Union and the state of Mecklenburg-West Pomerania (grant number ESF/14-BM-A55-0046/16; PePPP). AH is supported by the Hermann und Lilly Schilling-Stiftung für medizinische Forschung im Stifterverband.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.scr.2020.102079>.

References

- Ferenci, P., Stremmel, W., Członkowska, A., Szalay, F., Viveiros, A., Stättermayer, A.F., Bruha, R., Houwen, R., Pop, T.L., Stauber, R., Gschwantler, M., Pfeifferberger, J., Yurdaydin, C., Aigner, E., Steindl-Munda, P., Dienes, H.-P., Zoller, H., Weiss, K.H., 2019. Age and sex but Not ATP7B genotype effectively influence the clinical phenotype of wilson disease. *Hepatology* 69, 1464–1476. <https://doi.org/10.1002/hep.30280>.
- Naumann, M., Pal, A., Goswami, A., Lojewski, X., Japtok, J., Vehlouw, A., Naujock, M., Günther, R., Jin, M., Stanslowsky, N., Reinhardt, P., Sternecker, J., Frickenhaus, M., Pan-Montojo, F., Storkebaum, E., Poser, I., Freischmidt, A., Weishaupt, J.H., Holzmann, K., Troost, D., Hermann, A., 2018. Impaired DNA damage response signaling by FUS-NLS mutations leads to neurodegeneration and FUS aggregate formation. *Nature Commun.* 9 (1), 335. <https://doi.org/10.1038/s41467-017-02299-1>.
- Peter, F., Trilck, M., Rabenstein, M., Rolfs, A., Frech, M.J., 2017. Dataset in support of the generation of Niemann-Pick disease Type C1 patient-specific iPSC cell lines carrying the novel NPC1 mutation c.1180T>C or the prevalent c.3182T>C mutation - Analysis of pluripotency and neuronal differentiation. *Data Brief* 12, 123–131. <https://doi.org/10.1016/j.dib.2017.03.042>.