SUPPLEMENTARY DATA

SAMHD1 prevents autoimmunity by maintaining genome stability

Stefanie Kretschmer[1], Christine Wolf[1], Nadja König[1], Wolfgang Staroske[2], Jochen Guck[2,3], Martin Häusler[4], Hella Luksch[1], Laura A. Nguyen[5], Baek Kim[5,6], Dimitra Alexopoulou[2,7], Andreas Dahl[2,7], Alexander Rapp[8], M. Cristina Cardoso[8], Anna Shevchenko[9], Min Ae Lee-Kirsch[1*]

[1] Department of Pediatrics, Medizinische Fakultät Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

[2] Biotechnology Center, Technische Universität Dresden, Dresden, Germany

[3] Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge, UK

[4] Department of Pediatrics, University Hospital, University of Aachen, Aachen, Germany

[5] Center for Drug Discovery, Department of Pediatrics, Emory University, Atlanta, Georgia USA

[6] College of Pharmacy, Kyung-Hee University, Seoul, South Korea

[7] Center for Regenerative Therapies, Technische Universität Dresden, Dresden, Germany

[8] Department of Biology, Technische Universität Darmstadt, Germany

[9] Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany

* Correspondence to:

Min Ae Lee-Kirsch, M.D.
Department of Pediatrics, Medizinische Fakultät Carl Gustav Carus
Technische Universität Dresden
Dresden, Germany
Fetscherstr. 74, 01307 Dresden, Germany
phone: +49-351-458 6878, fax: +49-351-458 6333
email: minae.lee-kirsch@uniklinikum-dresden.de

Supplementary methods

Immunoprecipitation, Western blot

Cells were lysed in RIPA buffer (50 mM Tris HCl, pH 7.4, 150 mM NaCl, 1 mM Na_2EDTA, 1 % Triton X-100, 1 mM Na_3VO_4 and 20 mM NaF) containing 2 U/ml DNase I (Invitrogen), protease and phosphatase inhibitors (Roche). Lysates were incubated with GFP-Trap_A beads (ChromoTek) or with anti-cyclin A-coupled agarose beads (abcam and BD Biosciences) and analyzed by Western blotting using the following antibodies: SAMHD1 (ProteinTech Group), CCNA and p21 (BD Biosciences), CCNB1, CCND1, CCNE, p16 and p-p53 (S15) (Santa Cruz), CDK1 (Cell Signaling), Alexa Fluor 594-phalloidin (Invitrogen), GFP (Roche) and β-actin (Sigma Aldrich). Anti-phospho-T592-SAMHD1-antibody was generated by immunization of mice with the phosphorylated peptide CAPLI(pT)PQKKE (aa 586-597). Immunoreactive signals were detected by chemiluminescence (Lumi-Light PLUS, Roche).

Mass spectrometry

GFP-SAMHD1 was purified on beads (GFP-Trap_A, ChromoTek) from transfected HEK293T cell lysates and separated by one-dimensional SDS-PAGE on a 4-12 % bis-Tris gel (NuPAGE; Invitrogen). After visualization with Coomassie staining the experimental lane was cut into 6-8 slices and each slice was separately in-gel digested with trypsin [1]. Retrieved tryptic peptides were analyzed by LC-MS/MS on an Ultimate3000 nanoLC systems (Dionex) interfaced on-line to an LTQ linear trap or LTQ Orbitrap XL hybrid mass spectrometer (Thermo Fisher Scientific) via the robotic nanoflow ion source TriVersaNanoMate (Advion BioSciences Ltd.) as described [2, 3]. Acquired MS/MS spectra were searched against Human IPI (July 2010) and full protein (NCBI, July 2011) databases using MASCOT software (Matrix Science, v.2.2.0) under the following settings: mass tolerance was ± 2 Da and ± 5 ppm for precursors for LTQ and LTQ Orbitrap XL, respectively, and ± 0.5 Da for fragments; variable modifications were propionamide (C), carbamidomethyl (C), and oxidation (M); and enzyme specificity was trypsin. Protein hits were evaluated using Scaffold software

(Proteome Software, v.3.1.4.1) under the following settings: minimal number of matched peptides - 2, peptide threshold – 95 %, protein threshold – 99 %. false discovery rate (FDR) calculated by Scaffold software was 0.1 % and 0.5 % for proteins and peptides, respectively.

Comet assay

Comet assays were performed using a modified version of the method described by Olive and Banath [4]. Frosted microscope slides (Fisher Scientific) were pre-coated with melted 0.5 % agarose (type II, Sigma) in PBS followed by 1 % agarose (type II, Sigma). 1.5 x 10^5 cells synchronized by 24 h serum starvation were suspended in 150 µl 1 % low-melting point agarose (type VII, Sigma) melted in PBS at 50 °C and dispersed on pre-warmed pre-coated slides. Experiments were run in triplicate for each experimental condition. After gelling, slides were incubated in alkaline lysis solution (2.5 M NaCl, 0.1 M EDTA, 10 mM Tris, 1% Triton X-100, pH 10) at 4 °C for 16-18 h. Slides were then washed 10 times for 10 min in PBS. Prior to electrophoresis, slides were equilibrated for 20 min in electrophoresis buffer. Electrophoresis was conducted in alkaline electrophoresis buffer (0.3 M NaOH, 1 mM Na2EDTA, pH 13) at 1 V/cm and 500 mA for 25 min. 200 µl Antifade with Sybr Gold (20 mM DABCO, 0.1 µl/ml Sybr Gold, 20 mM Tris-HCl, 80 % glycerol) was added to each slide and covered with cover slips. Microscopy was carried out at 10x magnification. Images were captured using constant exposure times and 20 cells per slide were analyzed using CASP (CASPLab). The fluorescence intensity of the comet tail relative to the head reflects the number of DNA breaks.

Confocal microscopy, immunofluorescence and UVC irradiation

Cells were grown on glass cover slips in 24-well cell culture plates. For co-localization analysis, HEK293T cells were co-transfected with GFP-cyclin A and mCherry-SAMHD1 using polyethylenimine, fixed with 4 % paraformaldehyde, mounted with VectaShield containing DAPI and analyzed by confocal microscopy. For staining of DNA double-strand breaks,

fibroblasts were fixed with 4 % paraformaldehyde, permeabilized with 0.5 % Triton-X-100 in PBS and probed with mouse-anti-γH2AX (Ser-139; clone 2F3; Novus Biologicals) and rabbit-anti-53BP1 (Novus Biologicals). Slides were then probed with Alexa Fluor 546-conjugated goat-anti-mouse-IgG and Alexa Fluor 488-conjugated goat-anti-rabbit-IgG (Molecular Probes), respectively, mounted in VectaShield Mounting Medium containing DAPI (Vector Labs). γH2AX/53BP1 double positive nuclear foci were counted in at least 30 cells per slide. Fibroblasts were irradiated with UVC (20 J/m^2) and incubated for 6 h prior to staining. Fluorescence images were captured using a LSM510 confocal microscope (Zeiss). The scatter plot was created with ZEN2012 (Zeiss).

RNA sequencing

For RNA sequencing mRNA was enriched from 2 µg total RNA using magnetic poly-dT bead based separation (Illumina). Strand-specific RNA sequencing libraries were prepared using the TruSeq Stranded mRNA Sample Prep Kit (Illumina). Libraries were checked for integrity by capillary electrophoresis (Fragment Analyzer, Advanced Analytical) and quantified by qPCR using the KAPA Library Quantification Kit (KAPA Biosystems). Equimolar amounts of barcoded libraries were pooled and sequenced using the HiSeq 2000 system (Illumina). On average 41 million 75 bp single end reads were generated for each sample with an average mappability of 90 %. For data analyses a splice junction library with a length of 120 nucleotides (60 + 60) per splice junction was created based on known exon-exon junctions according to the Ensembl Genes annotation (v. 67, May 2012) [5]. Alignment of the reads to the hg19 transcriptome was performed with pBWA [6]. A counts-per-gene table was created based on the overlap of the uniquely mapped reads with the Ensembl Genes annotation for hg19, using BEDtools (v. 2.11) [7]. RPKM (reads per kilobase per million reads) values were calculated based on the raw read counts and used as a measure of absolute expression levels. Heat maps were constructed using GENE-E (Broad Institute) [8].

Fluorescence cross-correlation spectroscopy

In dual color FCCS, the excitation of two spectrally different dyes in the same detection volume allows measuring the interactions between differently labeled molecules [9]. Two superimposed laser lines of different wavelength are focused by a single microscope objective in order to generate overlapping focal volume elements. In this way, two spectrally distinct dyes can be simultaneously excited and their fluorescence emission detected via the same objective in two separate channels. In fluid environments, diffusion of fluorescent molecules through the focal spot induces fluorescence fluctuation. Both fluctuating fluorescence signals are subjected to auto-correlation analysis. Additionally, the calculated cross-correlation function is obtained by correlating intensity fluctuations of both channels with each other. Only when a dual labeled molecule translocates through the detection volume it contributes to the cross-correlation curve. Hence, the amplitude of the cross-correlation curve is directly proportional to the amount of double labeled species and allows a quantitative analysis of interactions.

1.5×10^4 HeLa cells were grown in fibronectin-coated 8-well chamber slides and transfected with 22 ng of GFP-cyclin A and 7.5 ng of mCherry-SAMHD1, respectively, using FuGene HD (Roche Diagnostics). FCCS was carried out on a commercial system consisting of a LSM780 and a ConfoCor3 (Zeiss). The 488 nm and 561 nm laser lines were attenuated by an acousto-optical tunable filter and directed via a 488/561 dichroic mirror onto the back aperture of a Zeiss C-Apochromat 40x, N.A. = 1.2, water immersion objective. The laser intensity in the focal plane was 0.67 kW/cm² (GFP) or 1.34 kW/cm² (mCherry), respectively. Fluorescence emission light was collected by the same objective and split into two spectral channels by a second dichroic (LP565). To remove any residual laser light, a 495-555 nm bandpass and 580 nm longpass emission filter, respectively, were employed. The fluorescence was recorded by avalanche photodiode detectors (APDs) in each channel. Out-of-plane fluorescence was rejected with a 35 µm pinhole. The fluorescence signals were software-correlated by ZEN software (Zeiss) and evaluated with MATLAB (Mathworks) by using weighted Marquardt nonlinear least-square fitting routine. For brightness analysis,

correction of the measured brightness of mCherry particles containing a dark fraction was carried out as previously described [10].

Supplementary references

1 Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat Protoc 2006;1:2856-60.

2 Shevchenko A, Roguev A, Schaft D, Buchanan L, Habermann B, Sakalar C, et al. Chromatin Central: towards the comparative proteome by accurate mapping of the yeast proteomic environment. Genome Biol 2008;9:R167.

3 Junqueira M, Spirin V, Santana BT, Waridel P, Surendranath V, Kryukov G, et al. Separating the wheat from the chaff: unbiased filtering of background tandem mass spectra improves protein identification. J Proteome Res 2008;7:3382-95.

4 Olive PL, Banath JP. The comet assay: a method to measure DNA damage in individual cells. Nat Protoc 2006;1:23-9.

5 Flicek P, Amode MR, Barrell D, Beal K, Brent S, Chen Y, et al. Ensembl 2011. Nucleic Acids Res 2011;39:D800-D806.

6 Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 2010;26:589-95.

7 Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010;26:841-2.

8 Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545-50.

9 Bacia K, Kim SA, Schwille P. Fluorescence cross-correlation spectroscopy in living cells. Nat Methods 2006;3:83-9.

10 Tungler V, Staroske W, Kind B, Dobrick M, Kretschmer S, Schmidt F, et al. Single-stranded nucleic acids promote SAMHD1 complex formation. J Mol Med 2013;91:759-70.

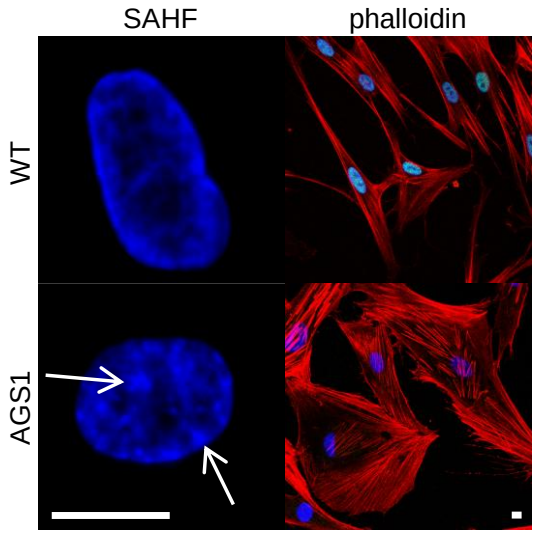
Supplementary table 1. Sequences of SAMHD1-siRNAs. All siRNAs including the control siRNA (Negative Control #1 siRNA) were purchased from Ambion.

name	sense (5´-3´)	antisense (5´-3´)
si-SAMHD1_1	GCAGAUAAGUGAACGAGAUtt	AUCUCGUUCACUUAUCUGCag
si-SAMHD1_2	GUAUCGCAUUUCUACAGCAtt	UGCUGUAGAAAUGCGAUACtt
si-SAMHD1_3	CGCAACUCUUUACACCGUAtt	UACGGUGUAAAGAGUUGCGag

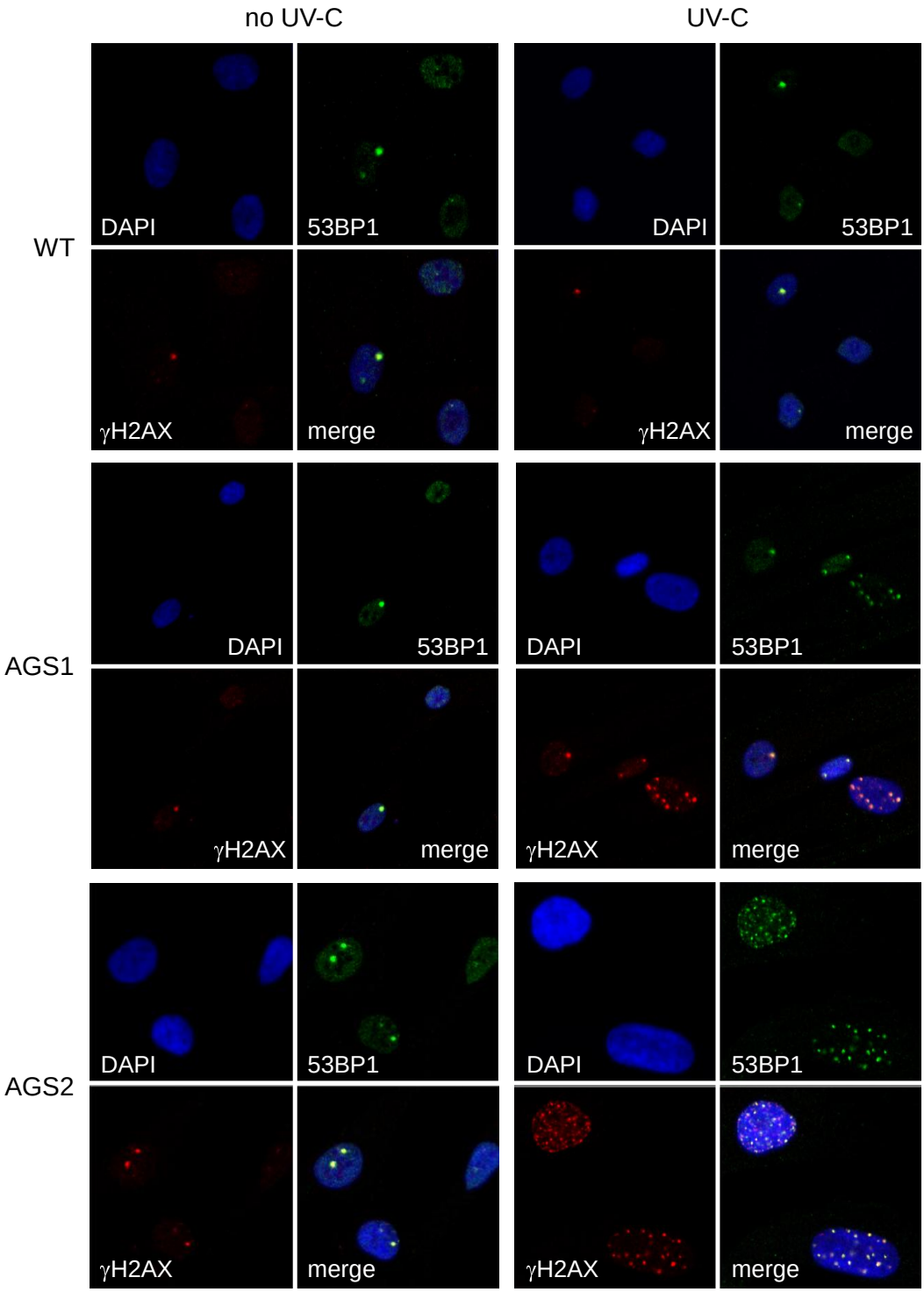
Supplementary table 2. Oligonucleotides used for quantitative RT-PCR. Oligonucleotides were purchased from Eurofins MWG Operon. IFI6 gene expression was determined using a pre-designed TaqMan Gene Expression Assay (Applied Biosystems).

name	**sequence (5´-3´)**
IFNB_F	ACCTCCGAAACTGAAGATCTCCTA
IFNB_R	TGCTGGTTGAAGAATGCTTGA
IFNB_P	FAM-CCTGTGCCTCTGGGACTGGACAATTG-TAMRA
IFI44_F	TCTATTCAATACTTCTCCTCTCAGATGATAG
IFI44_R	TGAGCAAAGCCACATGTACCA
IFI44_P	FAM-CCAGCGTTTACCAACTCCCTTCGAATTCTT-TAMRA
IFI27_F	AGCAGTGACCAGTGTGGCCAAAGT
IFI27_R	CTCCAATCACAACTGTAGCAATCC
IFI27_P	FAM-CCTCTGGCTCTGCCGTAGTTTTGCC-TAMRA
DDX58_F	AAAGCCTTGGCATGTTACACA
DDX58_R	TGGTCTACTCACAAAGCATTCCT
DDX58_P	FAM-AAGCATCTCCAAGCACAGTGTAATGGCA-TAMRA
GAPDH_F	GAAGGTGAAGGTCGGAGTC
GAPDH_R	GAAGATGGTGATGGGATTTC
GAPDH_P	FAM-CAAGCTTCCCGTTCTCAGCC-TAMRA

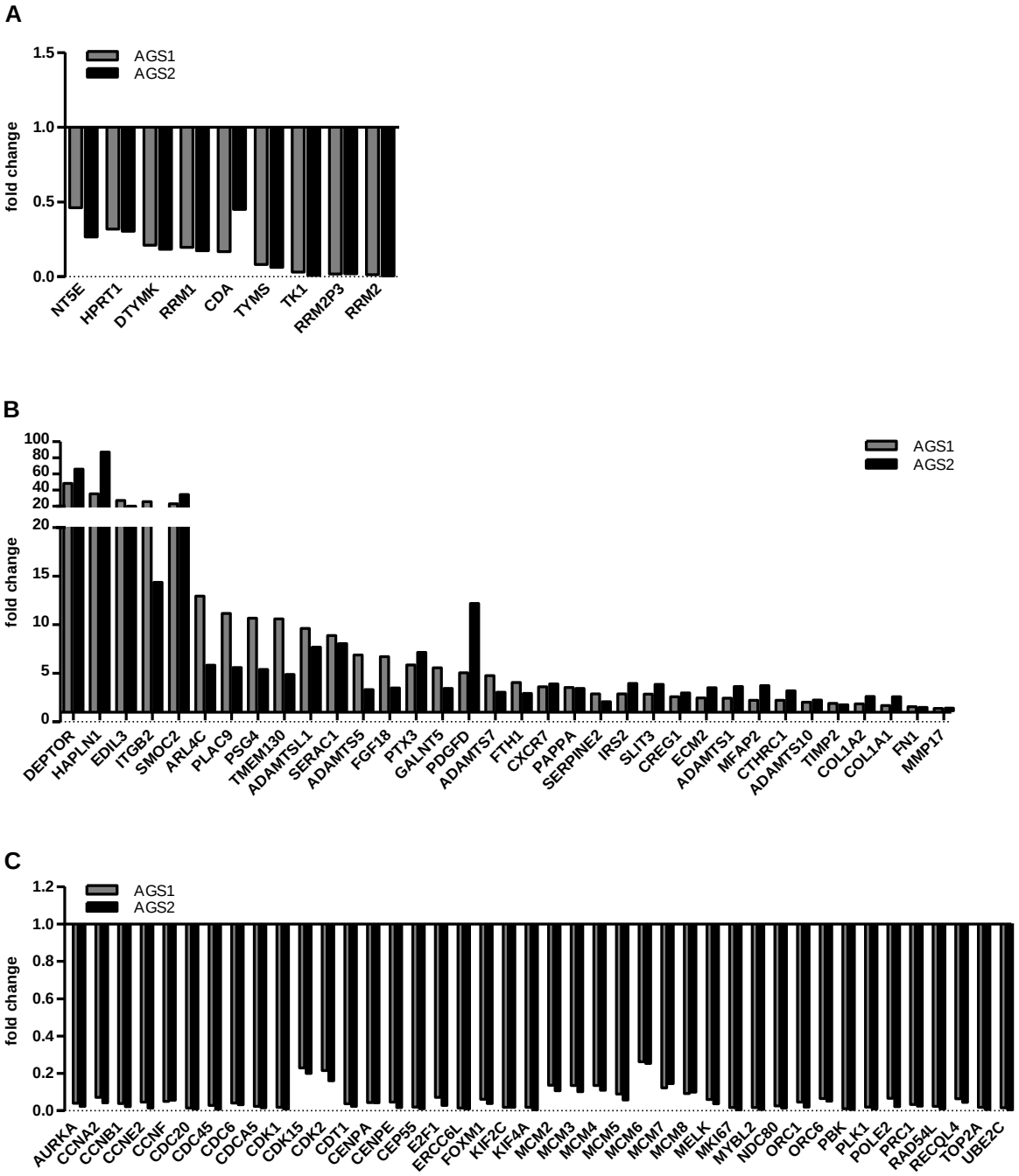
Supplementary figure 1. Senescence-associated morphological changes in SAMHD1-deficient fibroblasts. Formation of senescence-associated heterochromatin foci (SAHF, arrows) indicating chromatin silencing in AGS1. Irregular and enlarged morphology of patient cells shown by staining with the cytoskeleton marker phalloidin. Scale bar, 10 µm.



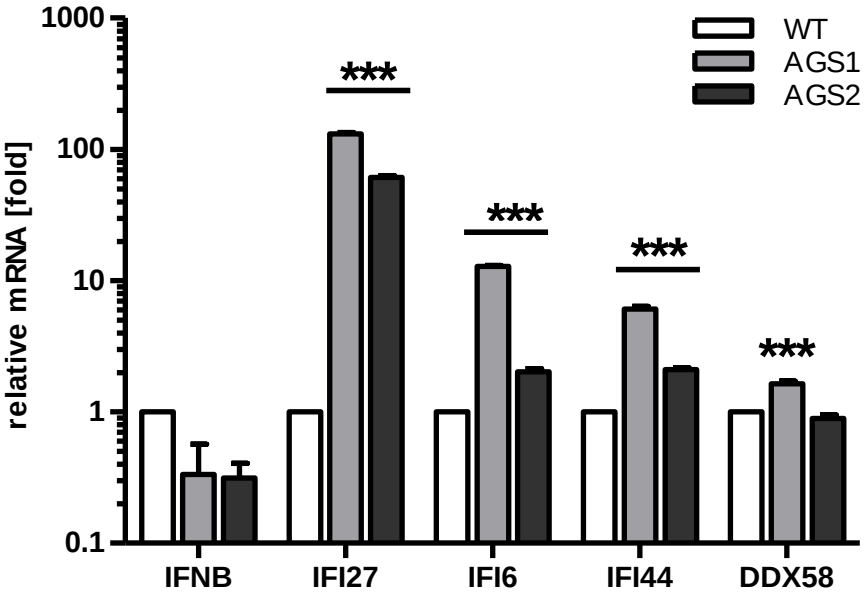
Supplementary figure 2. Immunofluorescence staining of DNA double-strand breaks in AGS patient fibroblasts. DNA double-strand breaks (DSBs) were visualized by co-staining for γH2AX (red) and 53BP1 (green) in native fibroblasts and 6 h after irradiation with low dose UV-C (20 J/m²). DSBs appear as double positive foci. Nuclei were counterstained with DAPI (blue). Shown are representative images from two independent experiments.



Supplementary figure 3. Expression of selected genes from transcriptome analysis. Shown are gene networks involved in dNTP synthesis (**A**), senescence-associated metabolic changes (**B**) and cell cycle progression (**C**). Indicated are the fold-changes of the RPKM values obtained by RNA sequencing analysis of patient fibroblasts (AGS1, AGS2) relative to the mean RPKM values of 4 wild type (WT) cell lines.



Supplementary figure 4. Validation of gene expression of selected IFN-stimulated genes in AGS fibroblasts. Gene expression of *IFNB*, *IFI27*, *IFI6*, *IFI44* and *DDX58* was determined by quantitative RT-PCR in fibroblasts of AGS patients (AGS1 and AGS2) and wild type control cells (WT). Indicated is the fold-change in gene expression relative to control. Gene expression was normalized to *GAPDH*. Data are represented as mean ± SEM of three technical replicates. ***: $P < 0.001$ by Student´s t test.



Supplementary figure 5. Mass spectrometry data showing phosphorylated SAMHD1. Mass spectrometric analysis of SAMHD1 digestion mixture detected two tryptic and three semi-tryptic peptides comprising phosphorylated threonine at position 592 (T592). The fragmentation spectrum of the doubly charged precursor ion with m/z 1081.552 which was confidently matched to the phosphorylated tryptic peptide sequence NFTKPQDGDVIAPLITpPQK is presented. The inset shows the low mass region of the spectrum with the phosphor-group at the T592 position.