Supplemental Data

Defective removal of ribonucleotides from DNA promotes systemic autoimmunity

Claudia Günther[1#], Barbara Kind[2#], Martin A.M. Reijns[3#], Nicole Berndt[1], Manuel Martinez-Bueno[4], Christine Wolf[2], Victoria Tüngler[2], Osvaldo Chara[5,6], Young Ae Lee[7], Norbert Hübner[7], Louise Bicknell[3], Sophia Blum[2], Claudia Krug[2], Franziska Schmidt[2], Stefanie Kretschmer[2], Sarah Koss[2], Katy R. Astell[3], Georgia Ramantani[8], Anja Bauerfeind[7], David L. Morris[9], Deborah S. Cunninghame Graham[9], Doryen Bubeck[10], Andrea Leitch[3], Stuart H. Ralston[11], Elizabeth A. Blackburn[12], Manfred Gahr[2], Torsten Witte[13], Timothy J. Vyse[9], Inga Melchers[14], Elisabeth Mangold[15,16], Markus M. Nöthen[15,16,17], Martin Aringer[18], Annegret Kuhn[19], Kirsten Lüthke[20], Leonore Unger[21], Annette Bley[22], Alice Lorenzi[23], John D. Isaacs[23], Dimitra Alexopoulou[24], Karsten Conrad[25], Andreas Dahl[24], Axel Roers[25], Marta E. Alarcon-Riquelme[4,26], Andrew P. Jackson[3], Min Ae Lee-Kirsch[2§]

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

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

[3] Medical Research Council Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK

[4] Centro de Genómica e Investigación Oncológica Pfizer-Universidad de Granada-Junta de Andalucía (GENYO), Granada, Spain

[5] Center for Information Services and High Performance Computing, Technical University Dresden, Dresden, Germany

[6] Institute of Physics of Liquids and Biological Systems, National University of La Plata, La Plata, Argentina

[7] Max Delbrück Centre for Molecular Medicine, Berlin-Buch, Germany

[8] Epilepsy Center, University of Freiburg, Freiburg, Germany

[9] Genetics & Molecular Medicine, King´s College London, London, UK

[10] Department of Life Sciences, Imperial College London, London, UK

[11] Rheumatic Diseases Unit, Centre for Genomic & Experimental Medicine, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK

[12] Centre for Translational and Chemical Biology, School of Biological Sciences, University of Edinburgh, Edinburgh, UK

[13] Hannover Medical School, Clinical Immunology, Hannover, Germany

[14] Clinical Research Unit for Rheumatology, University Medical Center, Freiburg, Germany

[15] Institute of Human Genetics, University of Bonn, Bonn, Germany

[16] German Center for Neurodegenerative Diseases (DZNE)

[17] Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany

[18] Rheumatology, Department of Internal Medicine III, Medizinische Fakultät Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

[19] Department of Dermatology, University Hospital, University of Münster, Münster, Germany

[20]Schwerpunktpraxis Rheumatologie, Dresden, Germany

[21] Städtisches Klinikum Dresden-Friedrichstadt, Dresden, Germany

[22] Department of Pediatrics, University of Hamburg, Germany

[23] Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, UK

[24] Center for Regenerative Therapies Dresden, DFG Cluster of Excellence, Technical University Dresden, Dresden, Germany

[25] Institute for Immunology, Medizinische Fakultät Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

[26] Arthritis and Clinical Immunology Program, Oklahoma Medical Research Foundation, Oklahoma City, USA

The authors declare no conflict of interests.

[#] Equal contribution

[§] Correspondence to Min Ae lee-Kirsch, Department of Pediatrics, Medizinische Fakultät Carl Gustav Carus, Technische Universität Dresden, Fetscherstr. 74, 01307 Dresden, Germany; phone: 0049-351-458 6887; email address: minae.lee-kirsch@uniklinikum-dresden.de.

Supplemental Methods

Sample collection used for replication of association. We used previously genotyped GWAS data from 5,478 individuals of European ancestry including 4,254 SLE patients and 1,224 controls genotyped using the Illumina HumanOmni1_Quad_v1-0_B chip and out-of-study controls obtained from three GWAS studies available at dbGaP with the appropriate approvals. . In order to obtain a quality-controlled working dataset satisfying current state-of-the-art criteria for association studies, data filtering was conducted using PLINK v1.07 1 (1) applying the following criteria: minimum total call rate per sample of 90%, minimum call rate per marker of 98%, minor allele frequency (MAF) threshold of 0.01%, Hardy-Weinberg Equilibrium (HWE) p-value for cases and controls at a minimum of 0.0001, and in addition at 0.01 only for controls, and finally a cutoff p-value of 0.00001 for differential missingness in no-call genotypes between cases and controls. To correct for stratification, principal component analysis was performed with smartpca, EIGENSOFT 4.0beta package 2 (2). This resulted in a λ_{GC} = 1.05 using the first 10 principal components. The final data set used for association analysis consisted of 4212 cases and 4065 controls.

Imputation and association analysis. Markers within the region of the three RNASEH2 genes (*RNASEH2A*, chr19:12804303-12813648; *RNASEH2B*, chr13:50909678-50970460; *RNASEH2C* chr11: 65717673-65720938; human genome assembly GRCh38) were selected for imputation with IMPUTE2 using the GWAS data of set 2 and the 1000 Genomes Project as reference (3). Prior to that, the three GWAS data sets obtained from the dbGaP database were pre-imputed with SHAPEIT using only EUR subpopulations as reference (4). For imputation, a restrictive QC-filter was applied (SNP genotyping rate ≥ 99%, sample genotyping rate ≥ 98%) without restriction of allele frequencies in order to include rare and low frequency variants. To ensure a high degree of reliable imputation, a conservative IMPUTE info_value threshold of > 0.8 for each marker was applied for association analysis at various minor allele frequency thresholds (<0.05, <0.01, <0.005, <0.001). To account for linkage disequilibrium and to avoid

the possible effect of redundancy, association was also tested including only markers with an info_value of > 0.8 and an LE r^2 < 0.5. Permutation was conducted to correct for multiple testing. Association was tested for each individual subunit and in an aggregate analysis assuming that all subunits form a single biochemical entity using the rare variant weighted aggregated statistic RWAS & LRT version 0.5 (5).

Comet assay. Comet assays were performed using a modified version of the method described by Olive et al. (6). 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. After gelling, slides were incubated in alkaline 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. For RNase H2 digestion, slides lysed in neutral lysis buffer were equilibrated 3 times for 20 min in RNase H2 reaction buffer (20 mM Tris-HCl, 10 mM $(NH_4)2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, 0.1% Triton X-100, pH 8.8). RNase H2 (New England Biolabs; 50 U per slide for murine cells, 100 U per slide for human cells) was then added in 50 µl of buffer, and slides were incubated for 16 h at 37 °C. Reactions were stopped by washing in 0.1% SDS in PBS for 10 min, followed by 3 washes for 10 min in PBS. Prior to electrophoresis, slides were equilibrated for 20 min in electrophoresis buffer. Electrophoresis was conducted in alkaline (0.3 M NaOH, 1 mM Na_2EDTA, pH 13) or neutral electrophoresis buffer (0.5 x TBE) at 1 V/cm and 500 mA for 25 min. Antifade with Sybr Gold (20 mM DABCO, 0.1 µl/ml Sybr Gold, 20 mM Tris-HCl, 80% glycerol; 200 µl) was added to each slide and covered with cover slips. Microscopy was carried out at 10-fold magnification. Images were captured using constant exposure times and 30 to 70 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. The Olive tail moment is defined as the product of the tail length and the fraction of total DNA in the tail.

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) 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). Alignment of the reads to the hg19 transcriptome was performed with pBWA (7). 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) (8). 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. Differential gene expression analysis based on the negative binomial distribution was carried out with DESeq R (v.1.10.1) (9). Pathway analysis was performed using Ingenuity Pathway Analysis (QIAgen). RNA sequencing data were deposited in the NCBI-GEO database.

Supplemental References

1. Purcell,S., Neale,B., Todd-Brown,K., Thomas,L., Ferreira,M.A., Bender,D., Maller,J., Sklar,P., de Bakker,P.I., Daly,M.J. et al 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**:559-575.
2. Patterson,N., Price,A.L., and Reich,D. 2006. Population structure and eigenanalysis. *PLoS. Genet.* **2**:e190.
3. Marchini,J., Howie,B., Myers,S., McVean,G., and Donnelly,P. 2007. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.* **39**:906-913.
4. Delaneau,O., Marchini,J., and Zagury,J.F. 2012. A linear complexity phasing method for thousands of genomes. *Nat. Methods* **9**:179-181.
5. Sul,J.H., Han,B., He,D., and Eskin,E. 2011. An optimal weighted aggregated association test for identification of rare variants involved in common diseases. *Genetics* **188**:181-188.
6. Olive,P.L., and Banath,J.P. 2006. The comet assay: a method to measure DNA damage in individual cells. *Nat. Protoc.* **1**:23-29.
7. Li,H., and Durbin,R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* **25**:1754-1760.
8. Quinlan,A.R., and Hall,I.M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* **26**:841-842.
9. Anders,S., and Huber,W. 2010. Differential expression analysis for sequence count data. *Genome Biol.* **11**:R106.
10. Ramantani,G., Kohlhase,J., Hertzberg,C., Innes,A.M., Engel,K., Hunger,S., Borozdin,W., Mah,J.K., Ungerath,K., Walkenhorst,H. et al 2010. Expanding the phenotypic spectrum of lupus erythematosus in Aicardi-Goutieres syndrome. *Arthritis Rheum.* **62**:1469-1477.
11. Sparks,J.L., Chon,H., Cerritelli,S.M., Kunkel,T.A., Johansson,E., Crouch,R.J., and Burgers,P.M. 2012. RNase H2-initiated ribonucleotide excision repair. *Mol. Cell* **47**:980-986.
12. Brzostek-Racine,S., Gordon,C., Van,S.S., and Reich,N.C. 2011. The DNA damage response induces IFN. *J. Immunol.* **187**:5336-5345.
13. Munoz-Fontela,C., Macip,S., Martinez-Sobrido,L., Brown,L., Ashour,J., Garcia-Sastre,A., Lee,S.W., and Aaronson,S.A. 2008. Transcriptional role of p53 in interferon-mediated antiviral immunity. *J. Exp. Med.* **205**:1929-1938.
14. Hasan,M., Koch,J., Rakheja,D., Pattnaik,A.K., Brugarolas,J., Dozmorov,I., Levine,B., Wakeland,E.K., Lee-Kirsch,M.A., and Yan,N. 2013. Trex1 regulates lysosomal biogenesis and interferon-independent activation of antiviral genes. *Nat. Immunol.* **14**:61-71.

Supplemental Table 1. Serologic findings in Aicardi-Goutières syndrome patients carrying mutations in *TREX1*, *RNASEH2A*, *RNASEH2B*, *RNASEH2C* or *SAMHD1* and their heterozygous parents. ANA: antinuclear antibodies; other autoantibodies: ss/dsDNA: single-stranded/double-stranded DNA; C3d-CIC, C1q-CIC: C3d/C1q circulating immune complexes; ENA: extractable nuclear antigens; C3/C4: complement C3/C4; ribP: ribosomal phosphoprotein; neg. serol.: negative serology. Normal values are indicated in square brackets. Clinical symptoms associated with autoimmune disease in AGS parents are given in round brackets. All AGS patients presented with signs of a leukencephalopathy, intracranial calcification, absence of findings of common prenatal infections, and/or a cerebral spinal fluid (CSF) white cell count of >5 white cells/mm^3 or raised levels of IFNα (>2 IU/l) in the CSF. [#]: Detailed clinical findings of these AGS patients have previously been described (10).

Patient Genotype	Patient Serology	Mother genotype	Mother Serology	Father genotype	Father Serology
TREX1 R114H hom	19 years ANA 1:1280 homogeneous/fine granular anti-nucleosome 138.36 U/ml [<30] anti-ssDNA 20.2 U/ml [<20] cardiolipin-IgG 15.5 U/ml [<10] C3d-CIC 49.24 µg/ml [<40]	R114H het	55 years ANA 1:2560 homogeneous/fine granular (arthralgia)	R114H het	50 years ANA 1:640 homogeneous/fine granular
TREX1 R114H hom	15 years anti-ssDNA 21.5 U/ml [<20] cardiolipin-IgG 15.97 U/ml [<10] C1q-CIC 65.49 µg/ml [<55] C3d-CIC 95.96 µg/ml [<40]				
TREX1[#] R114H + P212fs	2 years neg. serol.	R114H het	25 years neg. serol.	P212fs het	29 years anti-C1q 28.47 U/ml [<20]
TREX1[#] E198K hom	19 years ANA 1:1280 granular ENA positive	E198K het	43 years C4 0.14 g/l [0.16-0.32] (arthralgia, photosensitivity, alopecia, depression)	E198K het	44 years ANA 1:160 homogeneous/fine granular
TREX1 V201D hom	10 years ANA 1:160 fine granular ENA positive cardiolipin-IgG 41.51 U/ml [<10] anti-dsDNA 77.62 IU/ml [<30]	V201D het	33 years ANA 1:160 fine granular	V201D het	36 years ANA 1:80 fine granular
RNASEH2A[#] R186W hom	2 years neg. serol.	R186W het	34 years ANA 1:320 homogeneous/fine granular	R186W het	45 years neg. serol.
RNASEH2B[#] A177T hom	1 year cardiolipin-IgM 19.23 U/ml [<10]	A177T het	24 years cardiolipin-IgM 19.48 U/ml [<10]	A177T het	25 years neg. serol.
RNASEH2B A177T hom	1 year cardiolipin-IgG 15.4 U/ml [<10]	A177T het	35 years ANA 1:160 fine granular	A177T het	35 years neg. serol.
RNASEH2B A177T + c.136+1delG	6 years not available	c.136+1delG het	43 years C4 0.11 g/l [0.16-0.32] C3d-CIC 42.7 µg/ml [<40] (arthralgia, oral ulcers, dry eyes)	A177T het	48 years neg. serol.
RNASEH2B c.136+1delG het	4 years neg. serol.	No mutation	29 years neg. serol.	c.136+1delG het	29 years ANA 1:160 fine granular (recurrent oral ulcers)
RNASEH2C[#] D39Y + D115fs	4 years anti-ssDNA 46.42 U/ml [<20]	D39Y het	35 years anti-ssDNA 44.49 U/ml [<20]	D115fs het	42 years rheumatic factor 16,2 IU/ml [<10] anti-ssDNA 32.02 U/ml [<20] C3d-CIC 46.13 µg/ml [<40]
RNASEH2C[#] R69W hom	4 years C3 0.79 g/l [0.84-1.67]	R69W het	28 years ANA 1:80 fine granular	R69W het	31 years neg. serol.
RNASEH2C P76L hom	12 years ANA 1:160 fine granular cardiolipin-IgG 16.2 U/ml [<10] anti-dsDNA 33.14 IU/ml [<30] anti-ssDNA 21.8 U/ml [<20] anti-nucleosome 69.27 U/ml [<30] anti-ribP 21.39 U/ml [<10]	P76L het	37 years unremarkable (arthralgia, photosensitivity, hypothyroidism)	P76L het	39 years ANA 1:160 fine granular anti-ssDNA 42.3 U/ml [<20]
RNASEH2C K90del het	1 year C1q-CIC 75.54 µg/ml [<55]	K90del het	32 years ANA 1:320 homogeneous/fine granular	No mutation	28 years neg. serol.
SAMHD1[#] H167Y hom	9 months neg. serol.	H167Y het	28 years cytoplasmic antigens 1:320 fine granular [<1:80]	H167Y het	29 years ANA 1:160 homogenous/fine granular anti-C1q 32.4 U/ml [<20]

Supplemental Table 2. Phenotypic data of SLE patients with rare RNase H2 variants. [1]: The 20 year old son of this patient carries the same mutation and was found to have ANA, lymphopenia, reduced C3 complement, vitiligo and cold sensitivity with livoid discoloration of the fingers. Raynaud, Raynaud´s syndrome; Sjögren, secondary Sjögren´s syndrome.

Patient	RNase H2 subunit	Nucleotide change	Amino acid change	sex	phenotype
1	A	c.223G>C + c.880G>A	E75Q + E294K	M	Malar rash, photosensitivity, arthritis, ANA, anti-Ro, lymphopenia, Sjögren
2	A	c.361G>A	A121T	F	Subacute cutaneous lupus, photosensitivity, ANA, anti-Ro, anti-La
3	A	c.533C>T	A178V	F	Malar rash, arthritis, ANA, anti-dsDNA, anti-Ro, anti-La, Sjögren, hypothyroidism
4	A	c.662A>G	K221R	F	Malar rash, discoid rash, arthritis, anti-C1q, anti-cardiolipin-IgM, diffuse alopecia
5	A	c.662A>G	K221R	M	Raynaud, arthritis, anti-phospholipid syndrome, anti-cardiolipin-IgG, reduced C3
6	A	c.730A>G	I244V	F	Malar rash, photosensitivity, arthritis, ANA
7	B	c.314A>C	D105A	M	Subacute cutaneous lupus, photosensitivity, ANA, anti-Ro, anti-La, anti-Sm, low complement, lymphopenia, proteinuria, erythema multiforme
8	B	c.28G>A	G10R	F	Malar rash, discoid rash, arthritis, ANA, reduced C3
9	B	c.285T>G	F95L	F	Malar rash, arthritis, ANA, anti-dsDNA
10	B	c.529G>A	A177T	F	Malar rash, arthritis, ANA, anti-dsDNA, epilepsy
11	B	c.529G>A	A177T	F	Malar rash, photosensitivity, arthritis, ANA
12	B	c.697A>C	K233Q	F	Malar rash, lupus rash, photosensitivity, ANA, anti-Ro, arthritis, 1 late abortion, 2 small-for-date babies, autoimmune thyroiditis
13	B	c.744A>C	K248N	F	Malar rash, arthritis, ANA, anti-dsDNA, autoimmunue hepatitis, autoimmune thyreoditis, pernicious anemia
14	B	c.859G>T	A287S	F	Discoid rash, photosensitivity, Raynaud, ANA, anti-dsDNA, reduced C3, lymphopenia
15	B	c.859G>T	A287S	F	Malar rash, photosensitivity, arthritis, ANA, anti-dsDNA, urticarial vasculitis
16	B	c.859G>T	A287S	F	ANA, anti-dsDNA, anti-Ro, anti-La, pleuritis, glomerulonephritis, hemolytic anemia, Sjögren
17	B	c.859G>T	A287S	F	Malar rash, arthritis, ANA, anti-dsDNA, anti-cardiolipin
18	C	c.268_270delAAG	K90del	F	Photosensitivity, Raynaud, arthritis, anti-cardiolipin-IgG, Sjögren
19	C	c.328_329delGAinsCC	E110P	F	Malar rash, arthritis, glomerulonephritis, ANA, anti-dsDNA, anti-U1snRNP, lymphopenia
20	C	c.348+1G>A[1]	splice site F116fs	F	Malar rash, lupus rash, chilblain lupus, photosensitivity, arthritis, ANA, anti-dsDNA, lymphopenia, episodes of fever, autoimmune thyroiditis, alopecia
21	C	c.468G>T	A156A splice site	F	Discoid rash, photosensitivity, Raynaud, serositis, ANA, anti-snRNP, anti-Ro, anti-phospholipid-IgG, anti-phospholipid-IgM, Sjögren

Supplemental Table 3. Functional consequences of RNase H2 sequence variants. Experimentally determined effects on recombinant enzyme activity, complex stability and nuclear localization are indicated, and the predicted overall functional impairment *in vivo* summarized in the final column. Mutations found in SLE patients only are shown in red; mutations found in both controls and patients in orange; mutations found in controls only in black. For weighted burden analysis by logistic regression functional impairment of each variant was classed as neutral (-), mild (● or ●●) or severe (●●● or ●●●●). n.d. = not done.

		Reduced enzyme activity[1]	Reduced complex stability[2]	Reduced nuclear localization[3]	Functional impairment[4]
RNASEH2A	E75Q + E294K	●●	●●	-	●●
	E84D	-	●●	-	●●
	A121T	-	-	n.d.	-
	K165R	-	●●	-	●●
	A178V	●●●●	●●	-	●●●●
	L202S	●●	-	-	●●
	L202S + D205E	●●	-	-	●●
	K221R	●	●	-	●
	R239C	●●●●	-	-	●●●●
	I244V	●●●	●	-	●●●
	G276R	●●	●	-	●●
	T297A	●●	●●	-	●●
RNASEH2B	G10R	●●●	●●	-	●●●
	F95L	●●●	●●●●	●	●●●●
	D105A	●●●●	●●	-	●●●●
	N152S	●●●	●●	-	●●●
	A177T	-	●●●	-	●●●
	K233Q	●●●●	●●	-	●●●●
	K248N	●●●●	●●	●●●●	●●●●
	S277N	●●●	●●	-	●●●
	A287S	●●●●	●	●	●●●●
RNASEH2C	E110P	-	●	-	●
	K90del	-	●	-	●
	c.348+1G>A	●●●●	no complex formed	●●●	●●●●
	A156A[5]	n.d.	n.d.	n.d.	●●●●

[1] Effect on *in vitro* RNase H2 enzyme activity: ● very weak effect (< 80% of wild type, not significant); ●● weak effect (< 75% of wild type, $P < 0.01$); ●●● moderate effect (< 50% of wild type, $P < 0.01$); ●●●● strong effect (< 25% of wild type, $P < 0.001$).

[2] Effect on complex stability measured as a reduction in ThermoFluor melting temperature (Tm) of the protein complex compared to wild type: ● very weak effect (0.3-0.5 °C reduction in Tm, $P < 0.001$), ●● weak effect (< 2 °C reduction in Tm, $P < 0.001$), ●●● moderate effect (2-5 °C reduction in Tm, $P < 0.001$), ●●●● strong effect (> 5 °C reduction in Tm, $P < 0.001$).

[3] Effect on nuclear localization: ● weak effect (extranuclear localization of RNASEH2A and C in 10-20% of cells, $P < 0.05$); ●●● strong effect (extranuclear localization of RNASEH2A and C in 80-100% of cells, $P < 0.001$); ●●●● very strong effect (complete absence of nuclear localization, $P < 0.001$).

[4] This prediction is based on aggregating *in vitro* enzymatic and protein stability data with subcellular localization studies, with the strongest effect in the three categories taken to predict the severity of the mutation on *in vivo* function.

[5] The *RNASEH2C*-A156A mutation like the c.348+1G>A mutation disrupts an essential splice site, and therefore a strong effect on functional impairment is inferred.

Supplemental Table 4. Results of the aggregated tests for rare and low frequency variants. Statistical analysis was run for each RNase H2 subunit separately (A, B and C) and for the 3 subunits together (ABC) as part of a single biochemical unit using the rare variant weighted aggregate statistic (RWAS). A set of MAF thresholds in descending order was applied to analyze the effect of rare and low frequency variants on aggregated association. Corrected *P* values were obtained by applying 10,000 permutations.

		Info value ≥ 0.8			Info value ≥ 0.8 and $r^2 < 0.5$		
subunit	**MAF threshold**	**number of markers**	***P***	***P* corrected**	**number of markers**	***P***	***P* corrected**
A	no	173	0.000000	0.000700	27	0.005911	0.016598
	0.05	153	0.000000	0.000800	23	0.005911	0.016598
	0.01	22	0.000043	0.005999	12	0.012787	0.032497
	0.005	15	0.000132	0.007299	9	0.016093	0.037696
	0.001	9	0.000128	0.007799	5	0.011094	0.025697
B	no	314	0.000000	0.023198	55	0.015208	0.027497
	0.05	127	0,000000	0.038596	34	0.016463	0.029997
	0.01	71	0,000000	0.023598	18	0.003728	0.007899
	0.005	52	0,000000	0.006699	15	0.003599	0.006199
	0.001	40	0,000000	0.005599	6	0.000747	0.000900
C	no	614	0.000000	0.009299	96	0.000988	0.034197
	0.05	244	0.000000	0.007599	64	0.000042	0.025497
	0.01	232	0.000000	0.007099	57	0.000670	0.030597
	0.005	226	0.000000	0.005599	51	0.001269	0.036896
	0.001	146	0.000000	0.007599	25	0.001347	0.022298
ABC	no	1101	0.000000	0.000200	178	0.000001	0.001100
	0.05	524	0.000000	0.000300	121	0.000001	0.000400
	0.01	326	0.000000	0.000800	88	0.000000	0.000500
	0.005	293	0.000000	0.000500	75	0.000002	0.000700
	0.001	195	0.000000	0.000400	36	0.000001	0.000200

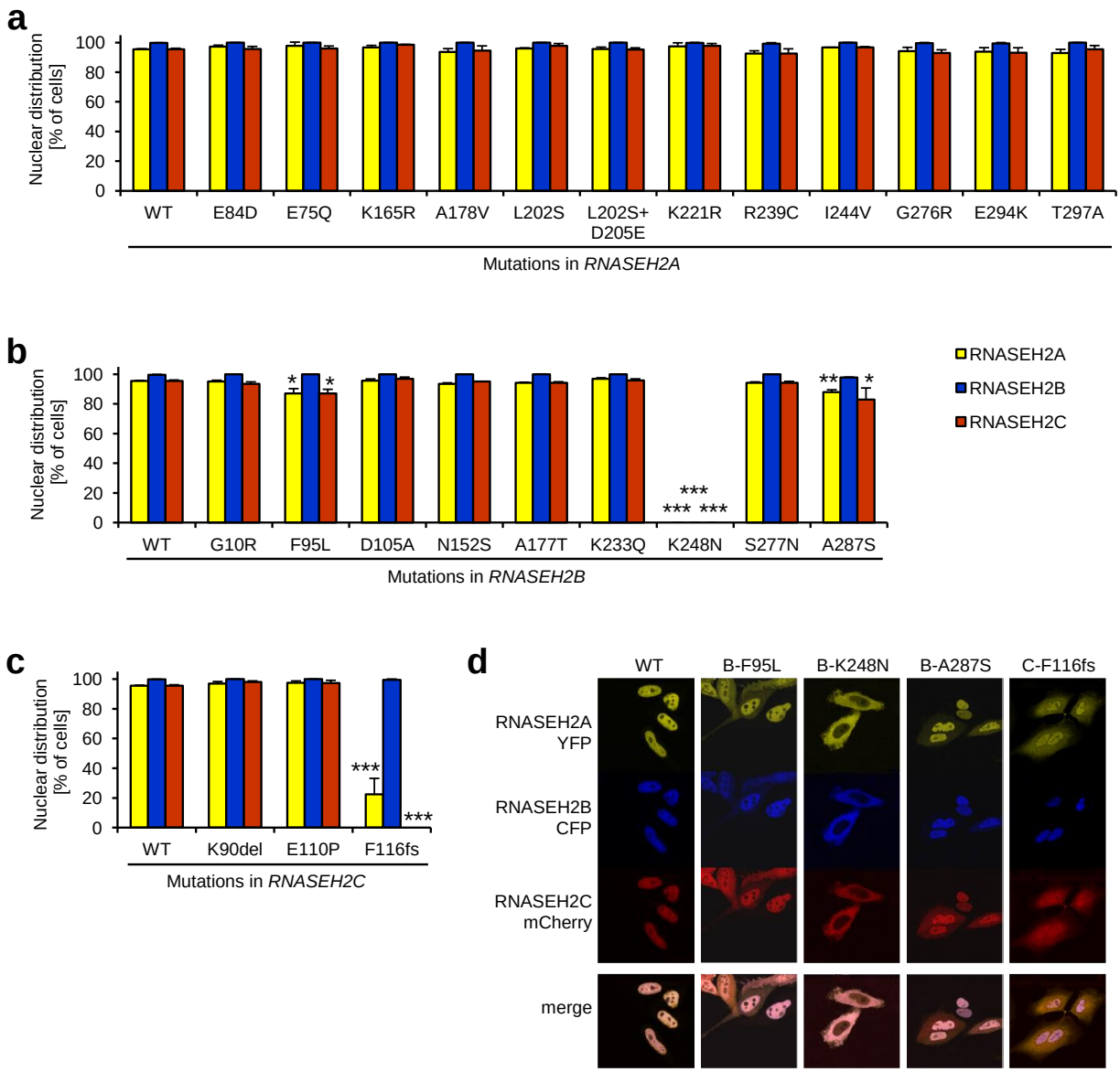
Supplemental Table 5. Primer sequences for mutation analysis of RNase H2 genes.

Gene	Name	Sequence (5´-3´)
RNASEH2A	RN2A-1F	GACCCGTCCTGCAGTATTA
	RN2A-1R	GAGGTTTGGAGACAGCACATC
	RN2A-2F	GGAAGGGATTCCTGGGTATG
	RN2A-2R	TCATGGCTCACTGCAGTCTC
	RN2A-3F	CGTGCTGGAGAAGAGGATTC
	RN2A-3R	CCCTGGCCATTTCAACATTA
	RN2A-4F	GGCAGGGAGCTTGAATGTTA
	RN2A-4R	CACTCCCACCTTGCTTCTGT
RNASEH2B	Dleu8-1bF	CCACCCGGAACAGACCCTTCT
	Dleu8-1bR	GGCCAGCTGTCCCAGAGAGAATC
	Dleu8-2F	AAGTAGAAGAAAGGAAAACAGG
	Dleu8-2R	AGCTATTCAGAAAGAGCAAGG
	Dleu8-3F	ATTCCCTCTTGATCTGATGGAGGATTTA
	Dleu8-3bR	AACCCGGTGACAGACCCACAAC
	Dleu8-4F	TGTTTTTCACTTTGAGATTAAGG
	Dleu8-4R	GATCATTAAAAGGAGAGAGAGG
	Dleu8-5F	CCCAGCCATGAGTTAATGTG
	Dleu8-5R	ACATTCAGTGTGTCTGTAAGC
	Dleu8-6F	GTTGATGACTCCTGAGGTAGCCACAT
	Dleu8-6cR	CAAAGCCTTCTTCTCATACTTGGTATCTCT
	Dleu8-7F	TAAATGGTCTGAAGGCCACC
	Dleu8-7R	ATGAGGCTTCTGTGATATTAAG
	Dleu8-8F	GGGTCAGAATTTGAATTCAGG
	Dleu8-8R	TCATAAACAGTGATATGATAAGC
	Dleu8-9F	GAAAAGTAGGACTTACTAAGTC
	Dleu8-9R	ACACCTTAAAATATATTAGAACC
	Dleu8-10F	GGTCTAAATTATACATGTTGGGT
	Dleu8-10R	AAACTGTAATCACAATCAAGCGT
	Dleu8-11bF	TAGTGGGGGATGCTTACCTTCCTTAA
	Dleu8-11bR	GTTTAAGAGAACATGAAATTGGCCTCTTC
RNASEH2C	AYP1-1F	GGAAAGCGTTAGGGGGTAAG
	AYP1-1R	CTCCTCCTCTTGGTCGTCAG
	AYP1-2F	TGGGATACGTGATGGTGACA
	AYP1-2R	GCTCCTTTATTGGGGTGATG

Supplemental Table 6. Primer sequences for RT-PCR. IFI6 gene expression was determined using a pre-designed TaqMan Gene Expression Assay (Applied Biosystems).

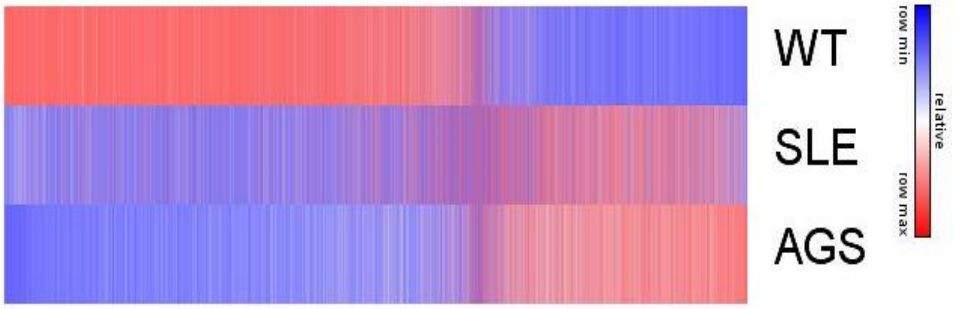
Target gene	Sequence (5´-3´)
IFNB_F	ACCTCCGAAACTGAAGATCTCCTA
IFNB_R	TGCTGGTTGAAGAATGCTTGA
IFNB_P	FAM-CCTGTGCCTCTGGGACTGGACAATTG-TAMRA
GAPDH_F	GAAGGTGAAGGTCGGAGTC
GAPDH_R	GAAGATGGTGATGGGATTTC
GAPDH_P	FAM-CAAGCTTCCCGTTCTCAGCC-TAMRA
IRF7_F	TCCCCACGCTATACCATCTACCT
IRF7_R	ACAGCCAGGGTTCCAGCTT
IRF7_P	FAM-AGCTGACAGGTCCTGCCCGAAGCC-TAMRA
HPRT1_F	TGGAAAGGGTGTTTATTCCTCAT
HPRT1_R	ATGTAATCCAGCAGGTCAGCAA
IFI27_F	TGCTCTCACCTCATCAGCAGT
IFI27_R	CACAACTCCTCCAATCACAACT
IFI44_F	CTGGGGCTGAGTGAGAAAGA
IFI44_R	AGCGATGGGGAATCAATGTA
IFI44L_F	ATCACCAGCATAACCGAGCG
IFI44L_R	ATCCATGCACAGTCCTGCTC
ISG15_F	GAGAGGCAGCGAACTCATCT
ISG15_R	CTTCAGCTCTGACACCGACA

Supplemental Figure 1. Effect of RNase H2 mutations on nuclear localization of the RNase H2 complex. HeLa cells were transfected with plasmids expressing each of the three subunits tagged at the N-terminus with YFP (RNASEH2A), CFP (RNASEH2B) or mCherry (RNASEH2C). The wild type heterotrimeric RNase H2 complex resides within the nucleus. (**a**-**c**) Rare variants identified by resequencing in each of the three subunits were expressed together with the two wild type subunits, and nuclear localization of each individual subunit was determined by examining ≥300 cells per mutation from three independent experiments. Shown are the means and SEM. **P* < 0.05; ***P* < 0.01; ****P* < 0.001 versus wild type by Fisher´s exact test. (**d**) Immunofluorescence images show reduced nuclear localization of RNase H2 for mutations in *RNASEH2B* (B-F95L, B-K248N, B-A287S) and *RNASEH2C* (C-F116fs, resulting from the c.358+1G>A mutation).

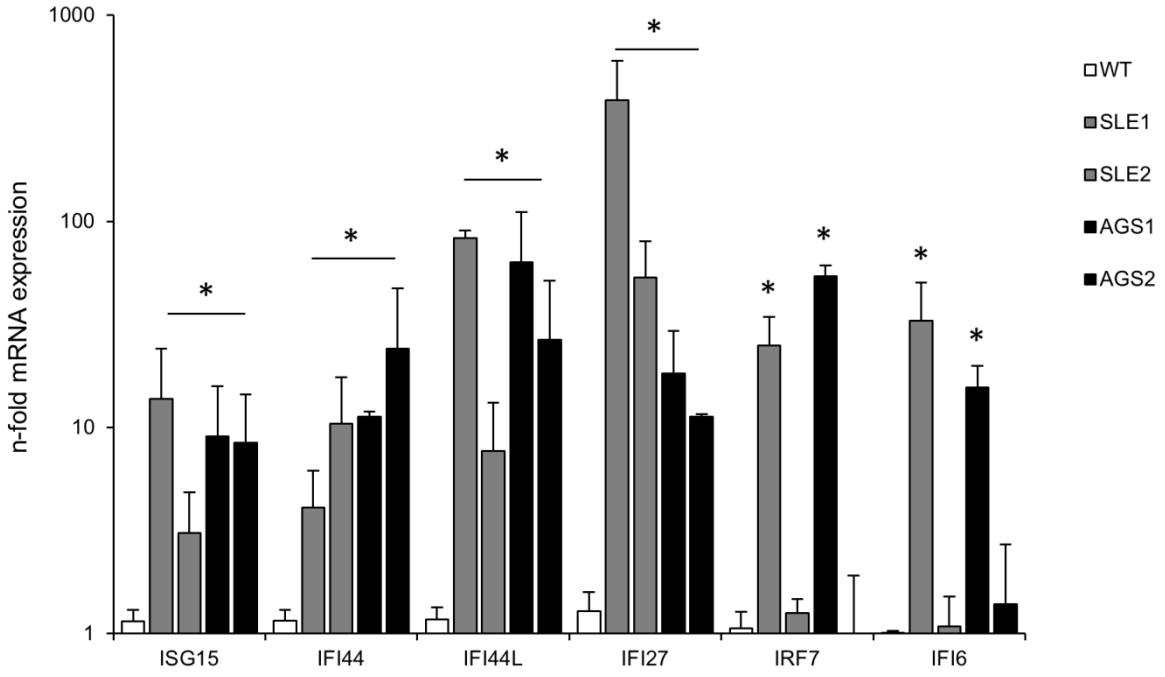


Supplemental Figure 2. Activation of p53-dependent and interferon-stimulated genes in RNase H2 mutant fibroblasts. (**a**) RNA sequencing analysis of fibroblasts from wild type controls (WT, n = 2) and patients with SLE (n =2) or AGS (n = 2). Heat map represents ~16.000 hierarchically clustered transcripts based on the mean RPKM values of each group. (**b**) Quantitative real-time RT-PCR of selected ISGs in fibroblasts of SLE and AGS patients compared to 3 wild type cell lines (WT). Error bars represent mean ± SEM. **P* < 0.05 versus wild type by Student´s *t* test.

a

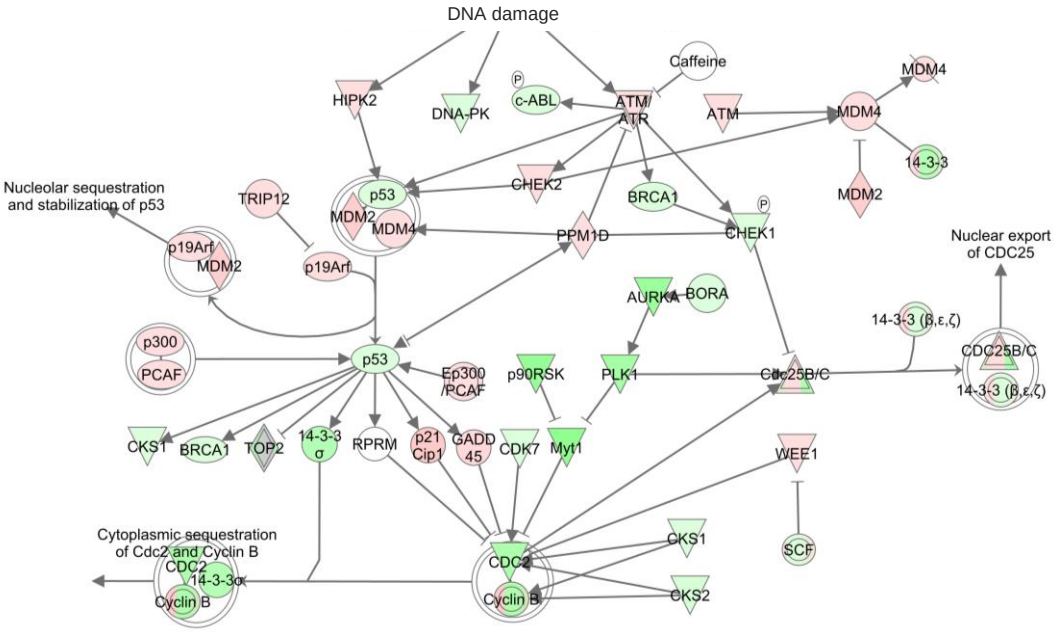


b

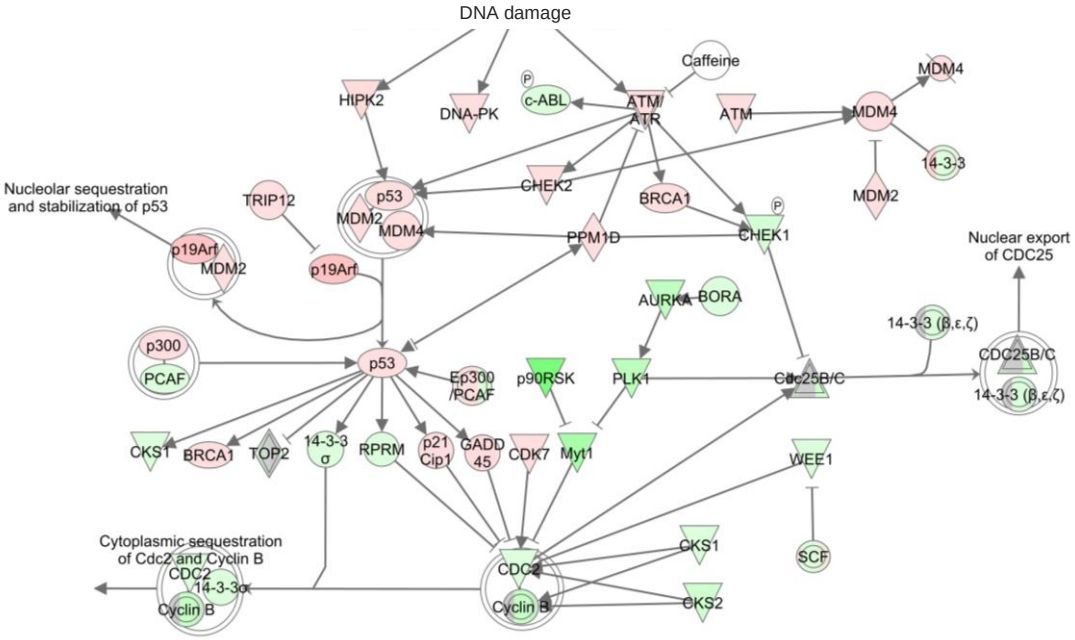


Supplemental Figure 3. Activation of DNA damage signaling in RNase H2-deficient fibroblasts. Differentially expressed genes obtained by RNA sequencing of patient fibroblasts (**a**, SLE; **b**, AGS) were mapped to the Ingenuity Pathway Analysis knowledge base and molecular networks were created using algorithmically generated pathways. The intensity of the node color indicates the degree of up-regulation (red) or down-regulation (green). Grey nodes indicate neither up- or down-regulated molecules. Mixed colored nodes represent a group of molecules.

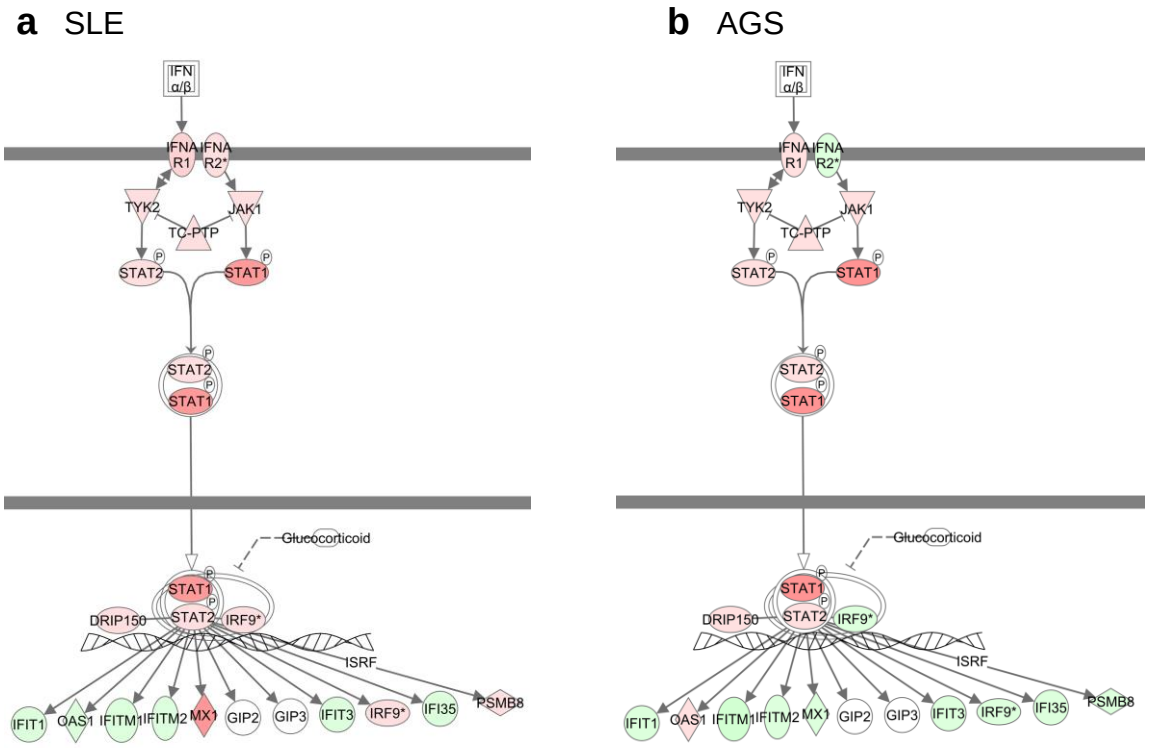
a SLE



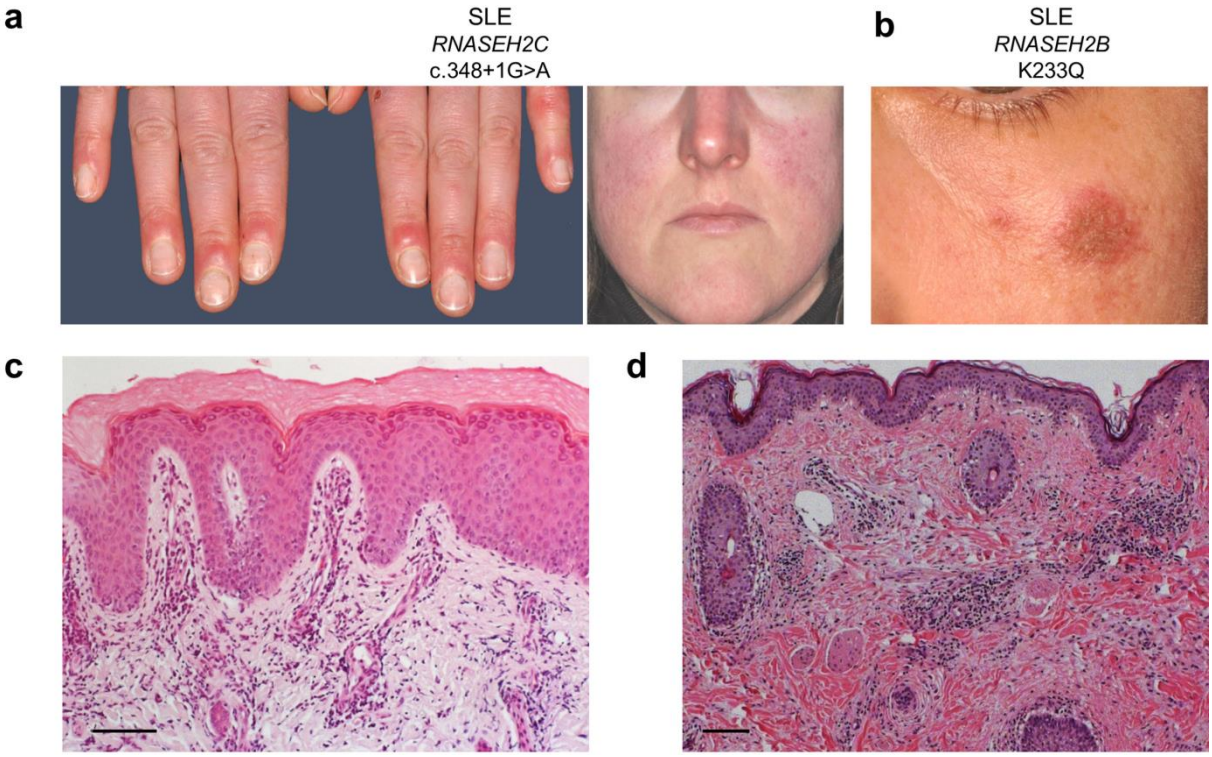
b AGS



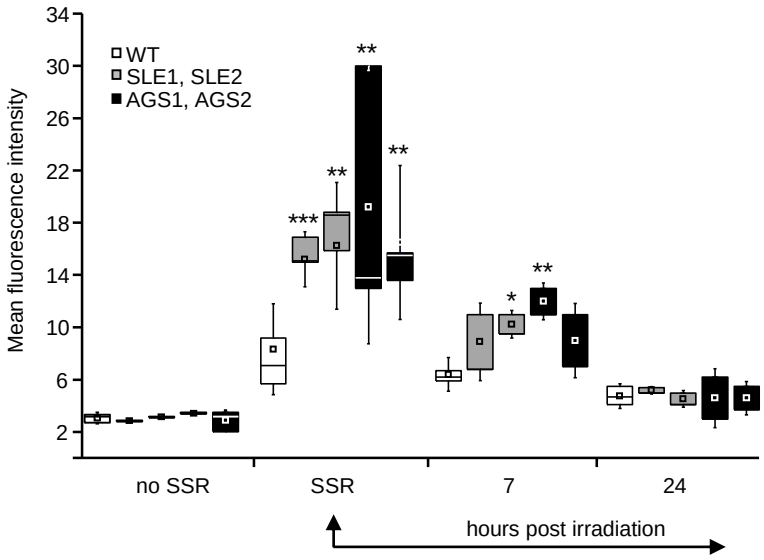
Supplemental Figure 4. Activation of type I-IFN signaling in RNase H2-deficient fibroblasts. Differentially expressed genes obtained by RNA sequencing of patient fibroblasts (a, SLE; b, AGS) were mapped to the Ingenuity Pathway Analysis knowledge base and molecular networks were created using algorithmically generated pathways. The intensity of the node color indicates the degree of up-regulation (red) or down-regulation (green). Grey nodes indicate neither up- or down-regulated molecules. Mixed colored nodes represent a group of molecules. Grey nodes indicate neither up- or down-regulated molecules. Mixed colored nodes represent a group of molecules.



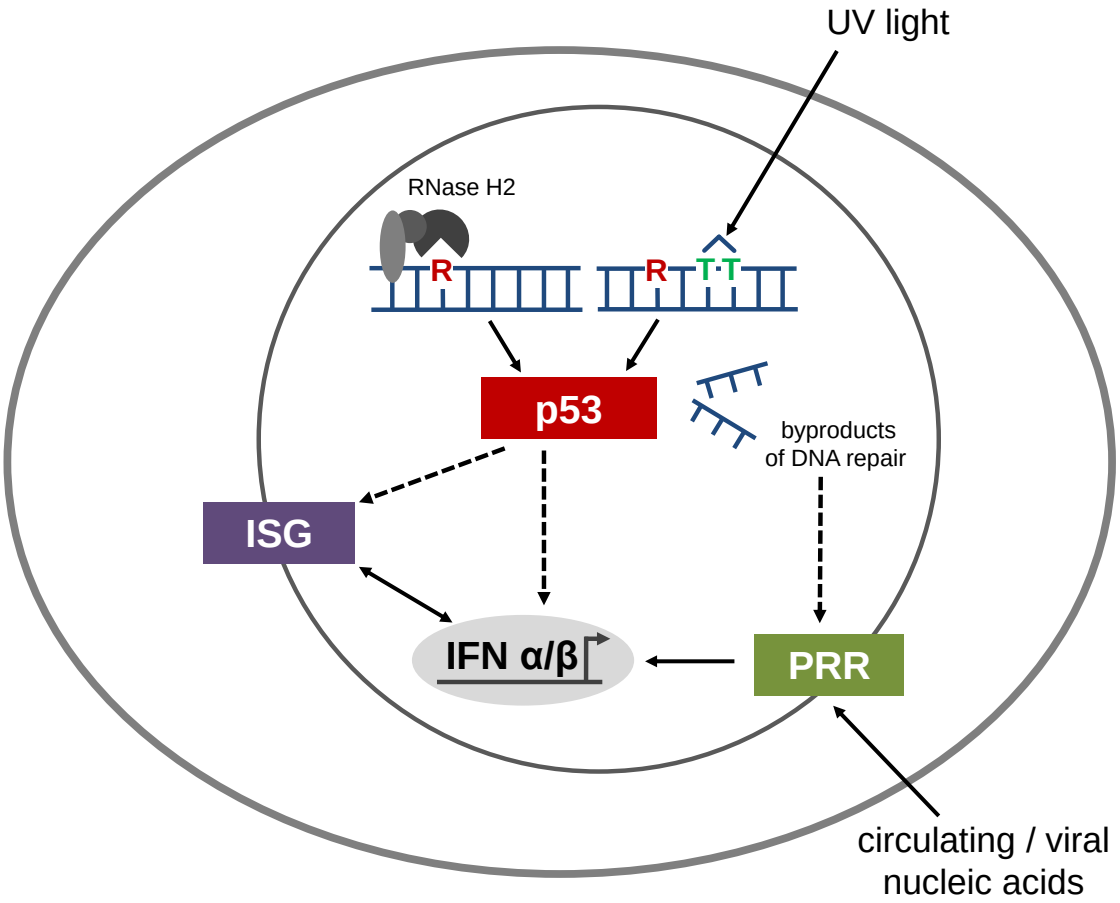
Supplemental Figure 5. Cutaneous phenotype of SLE patients carrying mutations in *RNASEH2C* (SLE1) and *RNASEH2B* (SLE2*).* (**a**) Patient SLE1 developed cold-induced livoid infiltrates on the dorsal side of her fingers primarily affecting the periungual area. Lesions partially ulcerated and resembled the typical picture of chilblain lupus also seen in patients with AGS. The facial appearance was characterized by erythematous macular exanthema with accentuation of the malar region and the bridge of the nose sparing the nasolabial fold. (**b**) Patient SLE2 had a history of malar rash and exanthema. After sun exposure she developed plaque-like erythematous infiltrates on her cheeks. These were long lasting and required sun protection and oral treatment with prednisone for healing. (**c**) A lesional biopsy from the finger of patient SLE1 shows perivascular and periadnexial lymphohistocytic infiltration along with acanthosis of the epidermis. Partially discrete interface dermatitis with alteration of the dermal epidermal junction is seen. (**d**) Histology of a lesional biopsy taken from the trunk of patient SLE2 shows superficial and deep perivascular as well as periadnexial lymphohistocytic infiltration with localized interphase dermatitis. (scale bar 100µm).



Supplemental Figure 6. Increased CPD formation in RNase H2-deficient fibroblasts exposed to solar simulated radiation. CPD formation in patient fibroblasts (SLE1, SLE2, AGS1, AGS2) and wild type cells (WT, n=4) after irradiation with 100 J/m^2 SSR. CPD formation was assessed by flow cytometry analysis immediately after irradiation and following 7 and 24 h of cultivation post-irradiation. Box plots depict the interquartile range (box), mean (square), median (line) and SD (whisker) of at least three independent experiments for each patient and for four independent wild type control cell lines. **P* < 0.05; ***P* < 0.01; ****P* < 0.001 versus wild type by *t* test.



Supplemental Figure 7. Model linking genomic ribonucleotide incorporation to abberant type I IFN production and induction of autoimmunity in AGS and SLE. Our data suggest deficient repair of misincorporated ribonucleotides leads to p53-dependent DNA damage signaling and enhanced cellular senescence. The heterotrimeric RNase H2 complex removes ribonucleotides from DNA by ribonucleotide excision repair (11). If left unrepaired, embedded ribonucleotides (indicated by an R) promote UV light-induced CPD formation (indicated by TT) due to conformational changes in the DNA that favour photocrosslinking, which may further fuel DNA damage signaling. Nucleic acid byproducts of DNA repair may bind and stimulate pattern recognition receptors (PRR), causing activation of innate immune signaling pathways and leading to type I IFN production. Alternatively, DNA damage signaling itself may sensitize cells to interferon-responsive stimuli, e.g. by p53-directed transcriptional upregulation of interferon-regulatory factors or by promoting transactivation of IFN-stimulated genes (ISGs) (12;13). Senescence-associated increased lysosomal activity may also sensitize cells to interferon-responsive stimuli (14). Enhanced type I IFN upregulation may be induced by PRR sensing of nucleic acids from viral infection or circulating nucleic acids released by apoptotic cells. Prolonged and enhanced type I IFN expression will favour the induction of autoimmune disease.



Supplemental Figure 8. UV-induced apoptosis in RNase H2-deficient fibroblasts. Patient fibroblasts (SLE1, SLE2, AGS1, AGS2) and wild type cells (WT, n=3) were irradiated with 20 J/m^2 UVC or 100 J/m^2 SSR, respectively. Sub$_{G1}$ fractions were determined 72 h post-irradiation by flow cytometry after propidium iodide staining. Shown are the means and SEM of five independent experiments.