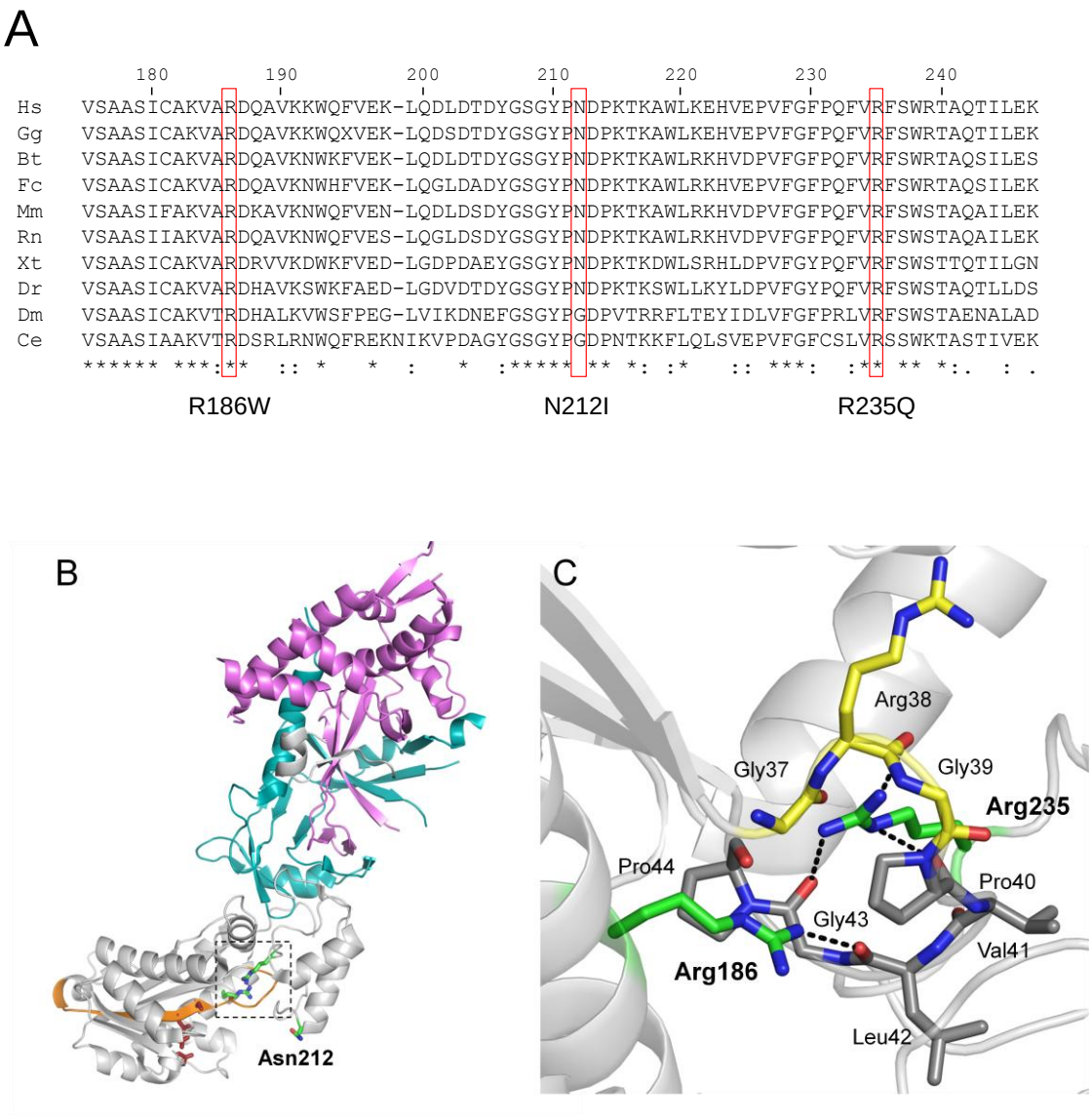
Supporting Information for the article:

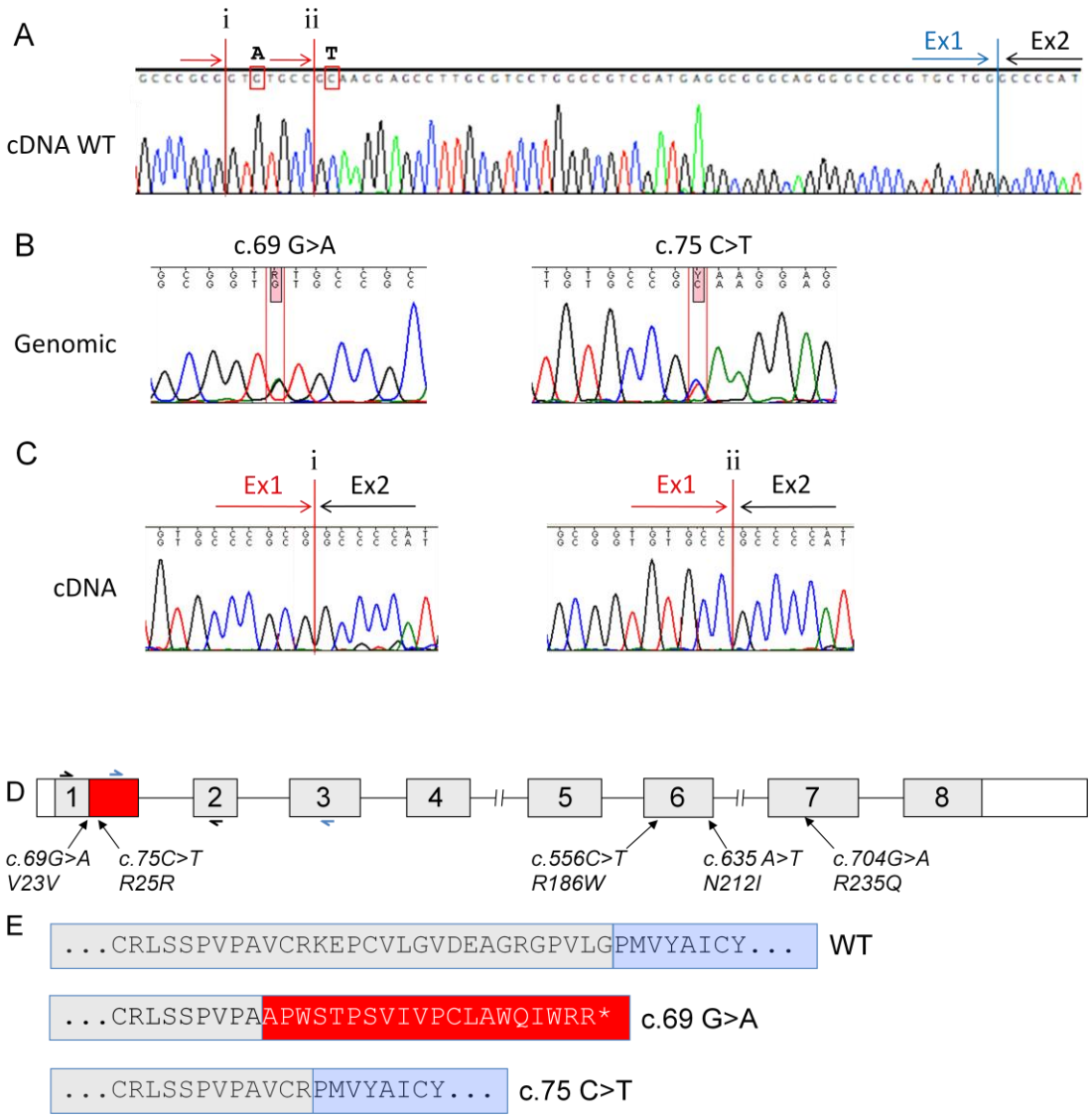
Synonymous mutations in *RNASEH2A* create cryptic splice sites impairing RNase H2 enzyme function in Aicardi-Goutières syndrome

Gillian I. Rice, Martin A.M. Reijns, Stephanie R. Coffin, Gabriella M.A. Forte, Beverley H. Anderson, Marcin Szynkiewicz, Hannah Gornall, David Gent, Andrea Leitch, Maria P. Botella, Elisa Fazzi, Blanca Gener, Lieven Lagae, Ivana Olivieri, Simona Orcesi, Kathryn J. Swoboda, Fred W. Perrino, Andrew P. Jackson, and Yanick J. Crow



Supp. Figure S1. Evolutionary conservation of RNASEH2A amino-acid residues affected by patient missense mutations, and their positions on the RNase H2 crystal structure. **A:** Evolutionary conservation of RNASEH2A residues affected in AGS patients (Arg186, Asn212, Arg235). Alignment constructed using Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/). "*" indicates positions which have a single, fully conserved residue; ":" indicates conservation between groups of strongly similar properties (scoring >0.5 in the Gonnet PAM250 matrix); "." indicates conservation between groups of weakly similar properties (scoring <0.5 in the Gonnet PAM250 matrix). Numbering indicates human RNASEH2A residues; red boxes mark residues mutated in AGS (R186, N212, R235). Species shown (accession number) are: Hs *Homo sapiens* (ENSG00000104889), Gg *Gorilla gorilla* (ENSGGOG00000005846), Bt *Bos taurus* (ENSBTAG00000009661), Fc *Felis catus* (ENSFCAG00000000474), Mm *Mus musculus* (ENSMUSG00000052926), Rn *Rattus norvegicus* (ENSRNOG00000003504), Xt *Xenopus tropicalis* (ENSXETG00000004762), Dr *Danio rerio*

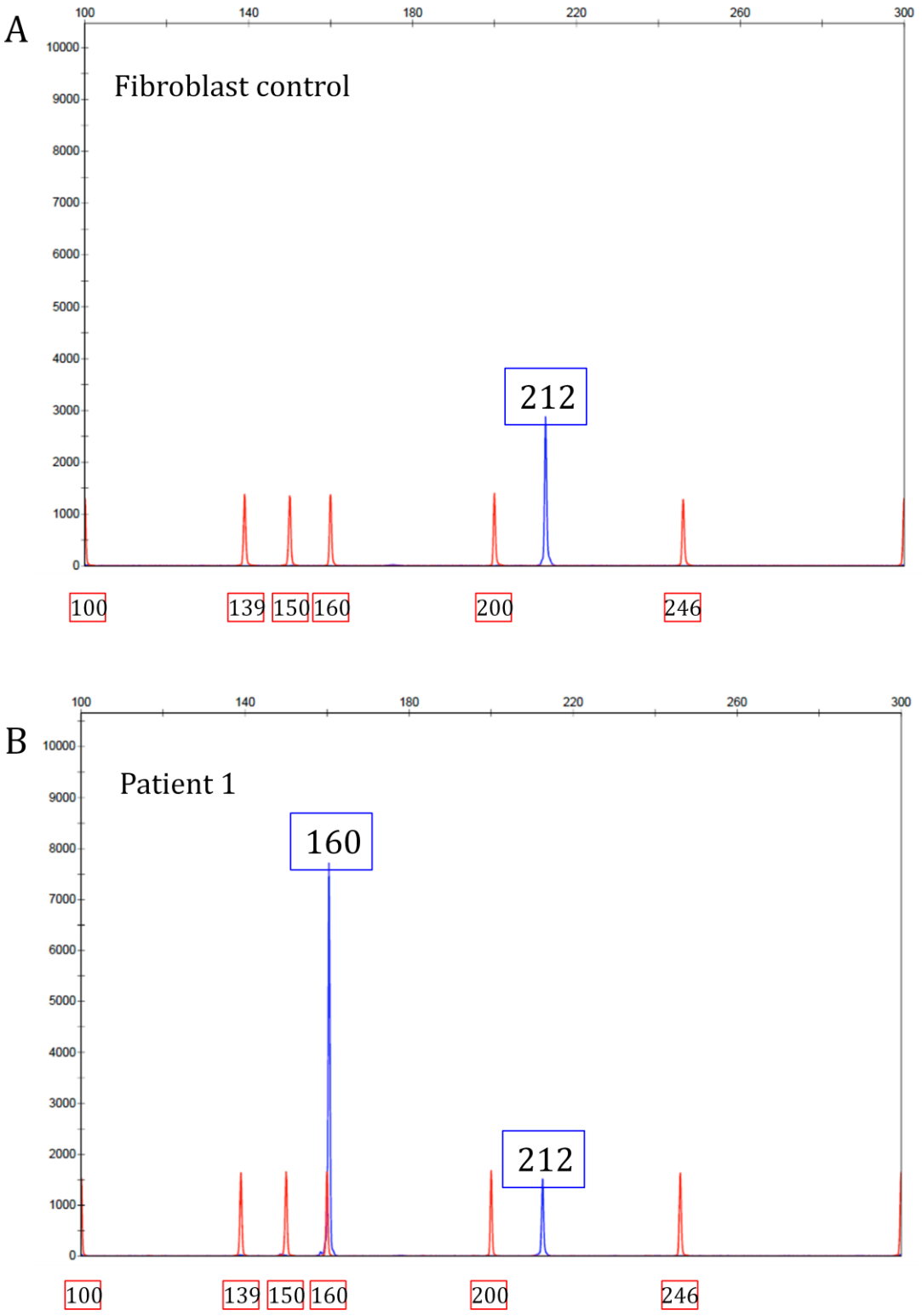
(ENSDARG00000018891), Dm *Drosophila melanogaster* (FBgn0031252), Ce *Caenorhabditis elegans* (T13H5.7). **B:** Structural roles of RNASEH2A arginine residues 186 and 235. The human RNase H2 heterotrimer (PDB ID: 3PUF) consists of the RNASEH2B (pink), RNASEH2C (cyan) and RNASEH2A (white) subunits. Side chains of residues affected in AGS patients (Arg186, Asn212, Arg235) are shown in green, side chains of catalytic site residues (Asp34, Glu35, Asp141, Asp169) in red. The portion of RNASEH2A deleted as a result of the c.75C>T (p.Arg25Arg) mutation is shown in orange. The square marks the region shown in **C:** Residues Arg235 and Arg186 (green) stabilize the turn between two anti-parallel β-sheets with direct hydrogen bonds (dotted lines) to the peptide backbone. Notably, one of these β-sheets contains two of the four proposed metal coordinating residues required for catalysis. Gly37, Arg38 and Gly39 (yellow) form the conserved GRG motif that is involved in specific interactions with the 2'-OH group of the RNA-DNA junction (Rychlik, et al., 2010).



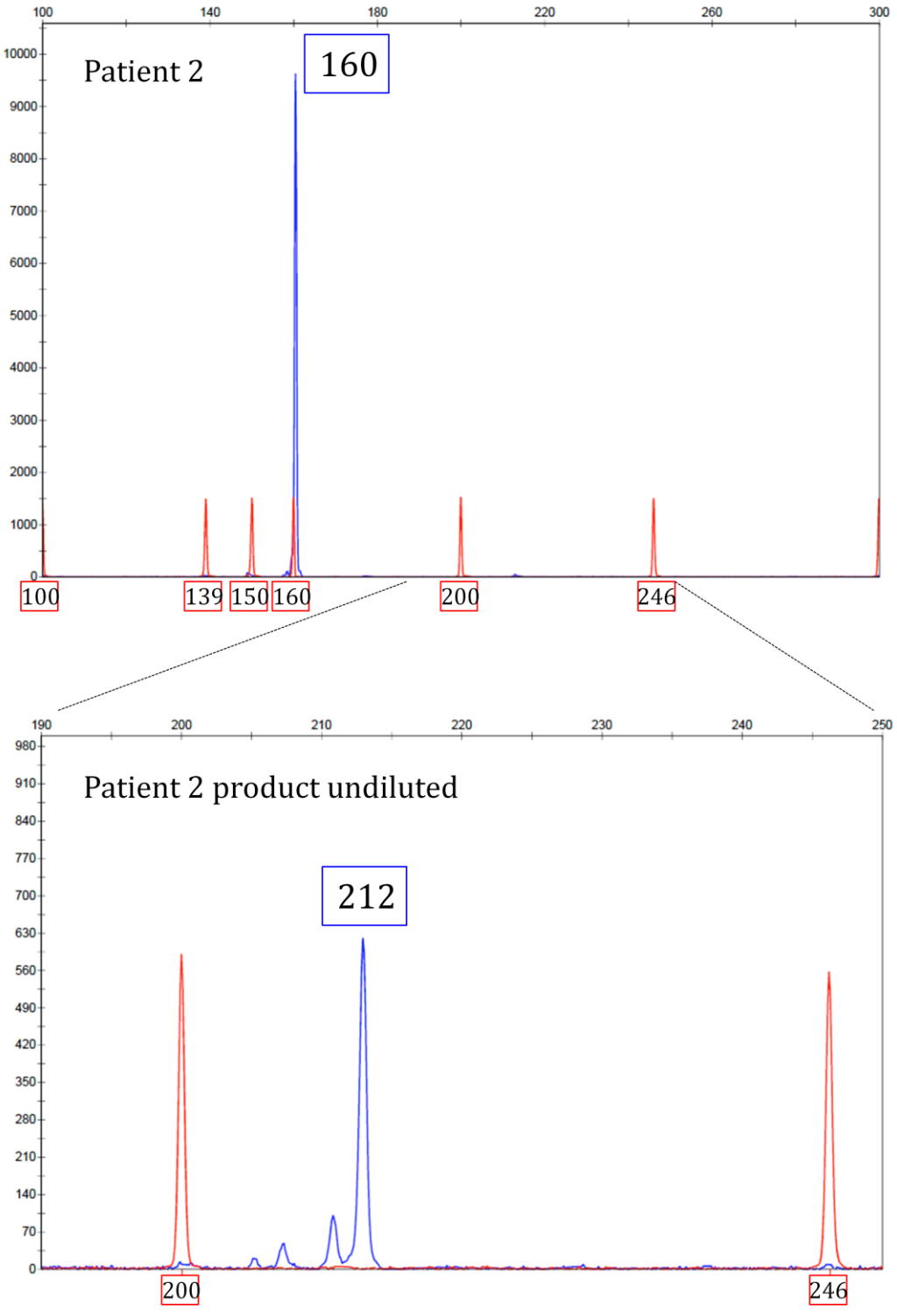
Supp. Figure S2. Synonymous *RNASEH2A* c.69G>A and c.75C>T mutations in AGS patients activate cryptic splice donor sites. **A:** cDNA full-length (WT) sequence trace showing the second half of exon 1 with positions of c.69G>A and c.75C>T highlighted by red boxes, and the normal exon 1/2 boundary shown by a blue vertical line and arrows. Red vertical lines show the position of the exon1/2 boundary resulting from the (i) c.69G>A and (ii) c.75C>T mutations. **B:** Genomic DNA sequence traces with c.69G>A and c.75C>T mutations in heterozygous individuals. Red boxes and red vertical lines highlight the mutated residue in each case. **C:** cDNA sequence traces of cDNA products that result from c.69G>A and c.75C>T mutations. Exon1/2 boundaries are marked with red vertical lines. Corresponding positions on the full-length sequence are shown in A with a red arrow, highlighting regions deleted in each case. **D:** Schematic of the *RNASEH2A* gene with AGS point mutations indicated. The red box highlights the region of exon 1 that is excluded during pre-mRNA splicing as a result of the c.75C>T mutation. Black arrow heads indicate Primer pair 1 (RNaseH2A1F and RNaseH2AdelR2;

Figure 1A); blue arrow heads indicate Primer pair 2 (RNaseH2AdelFnew and RNaseH2AdelRnew; Figure 1A). **E:** Schematic showing parts of the RNase H2A amino acid sequence for the full-length protein (WT), and those resulting from c.69G>A and c.75C>T mutations.

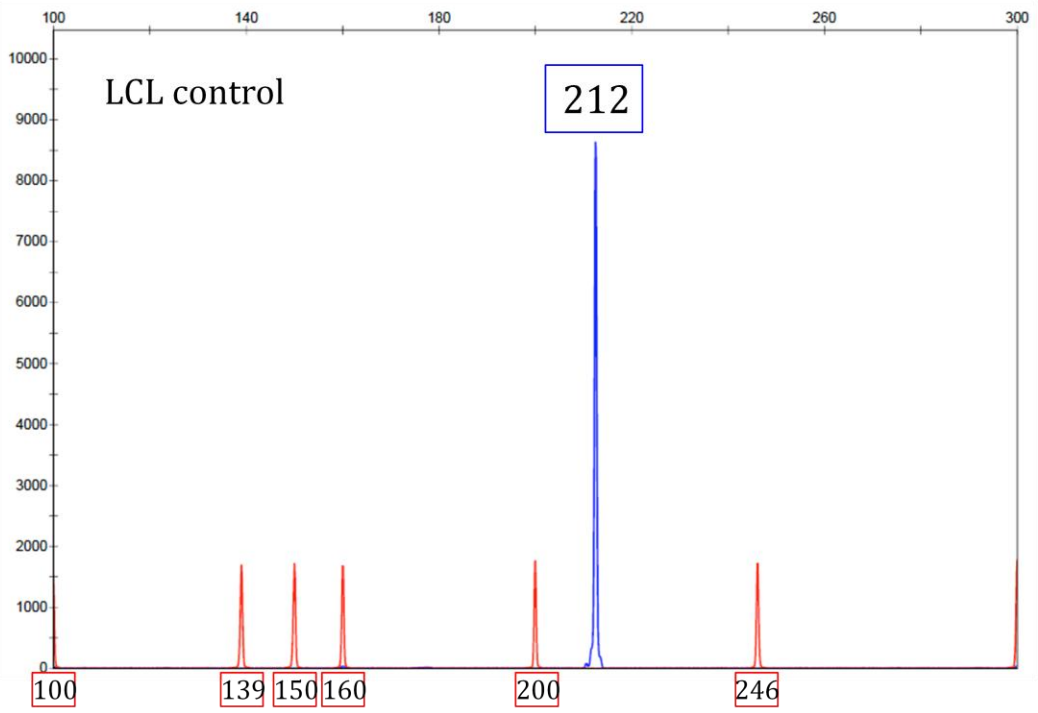
Supp. Figure S3



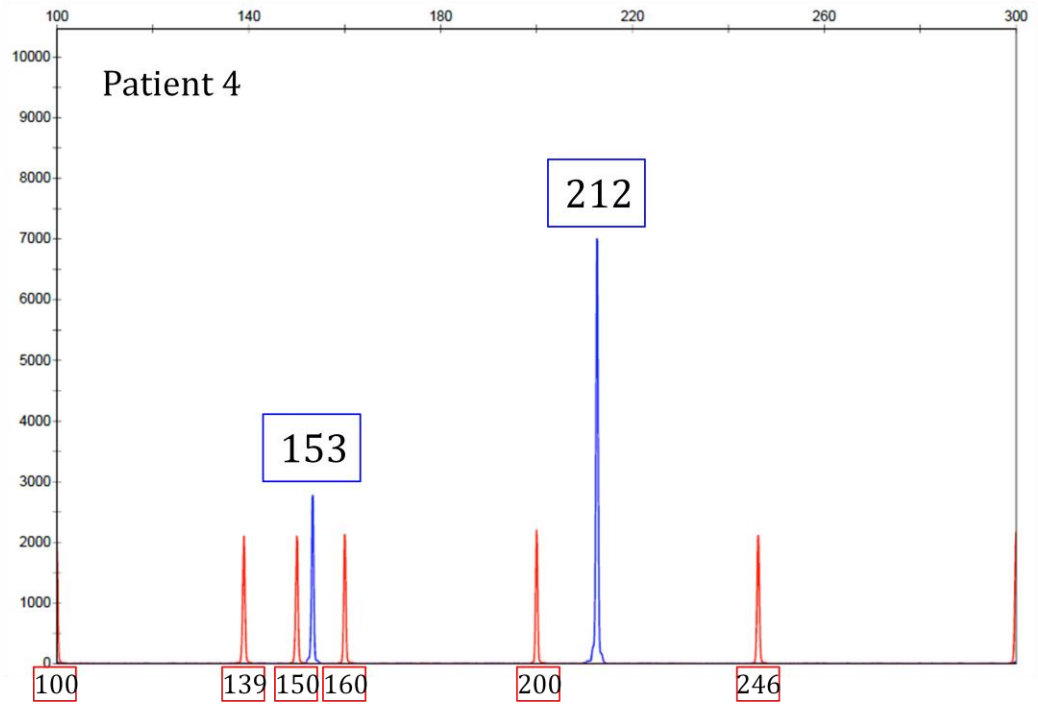
C


Patient 2
160
100
139
150
160
200
246
Patient 2 product undiluted
212
200
246

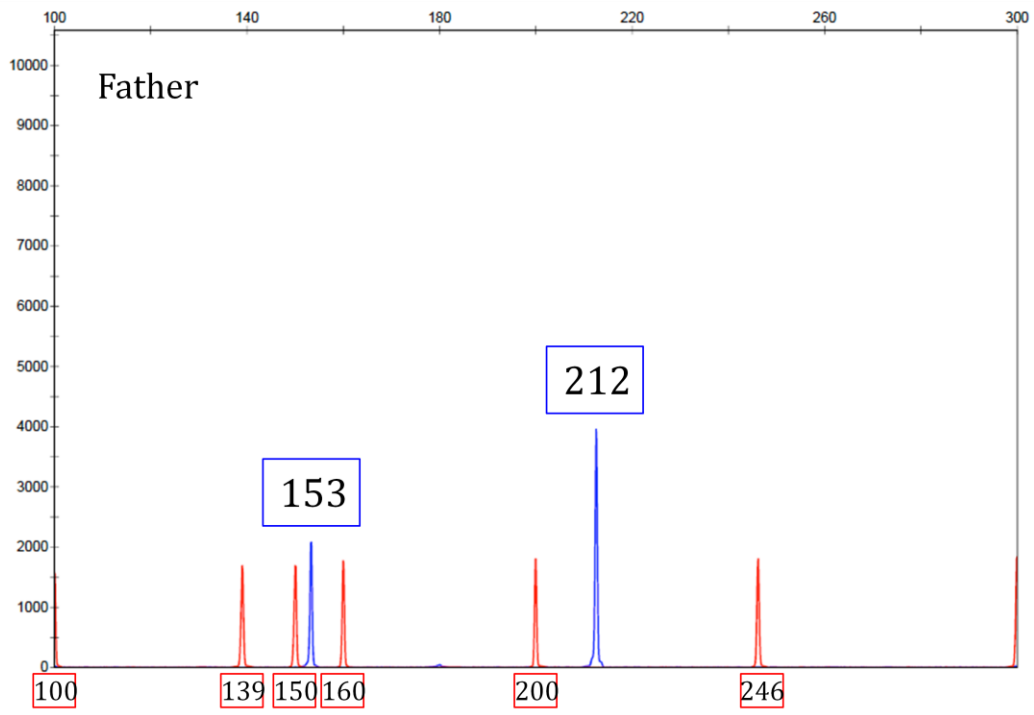
D


LCL control
212
100
139
150
160
200
246

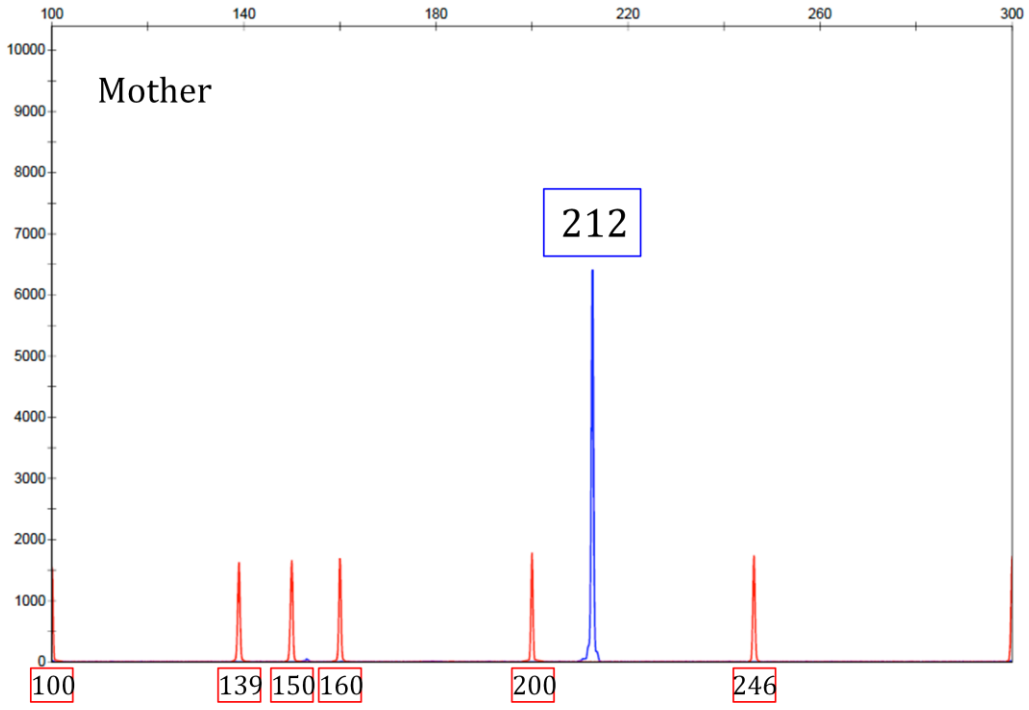
E


Patient 4
212
153
100
139
150
160
200
246

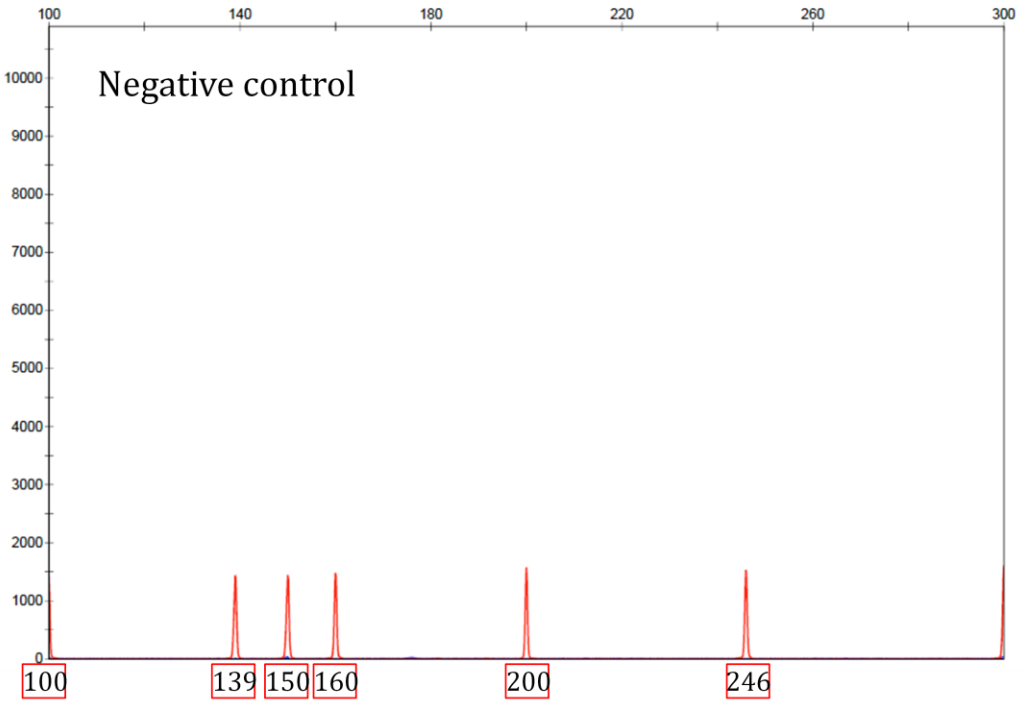
F


100
140
180
220
260
300
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
Father
212
153
100
139
150
160
200
246

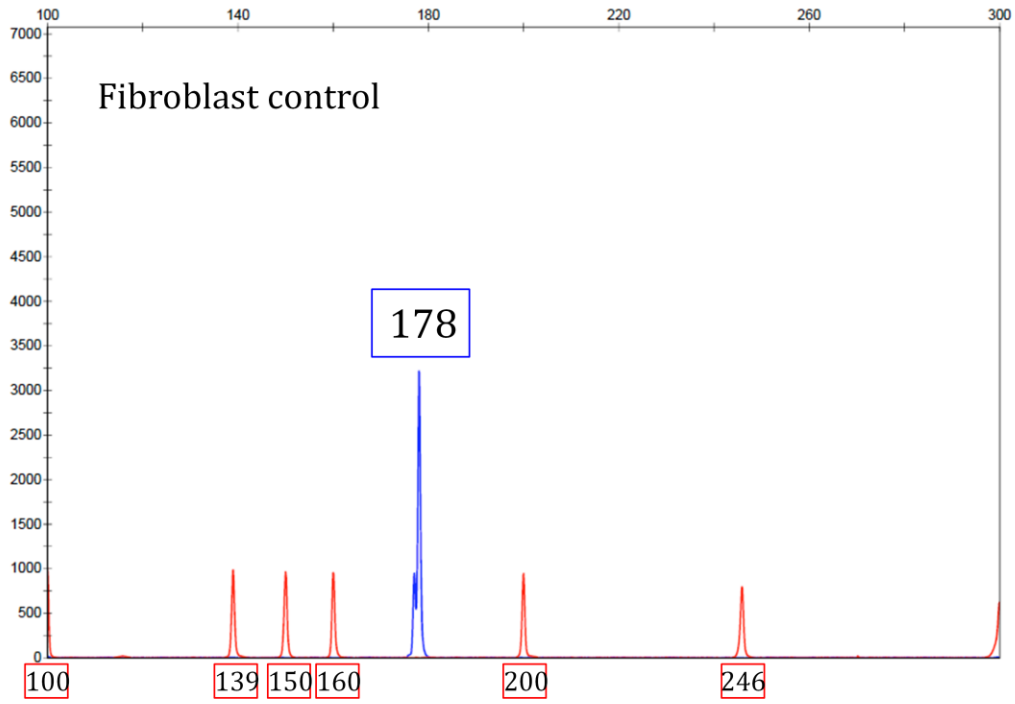
G


100
140
180
220
260
300
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
Mother
212
100
139
150
160
200
246

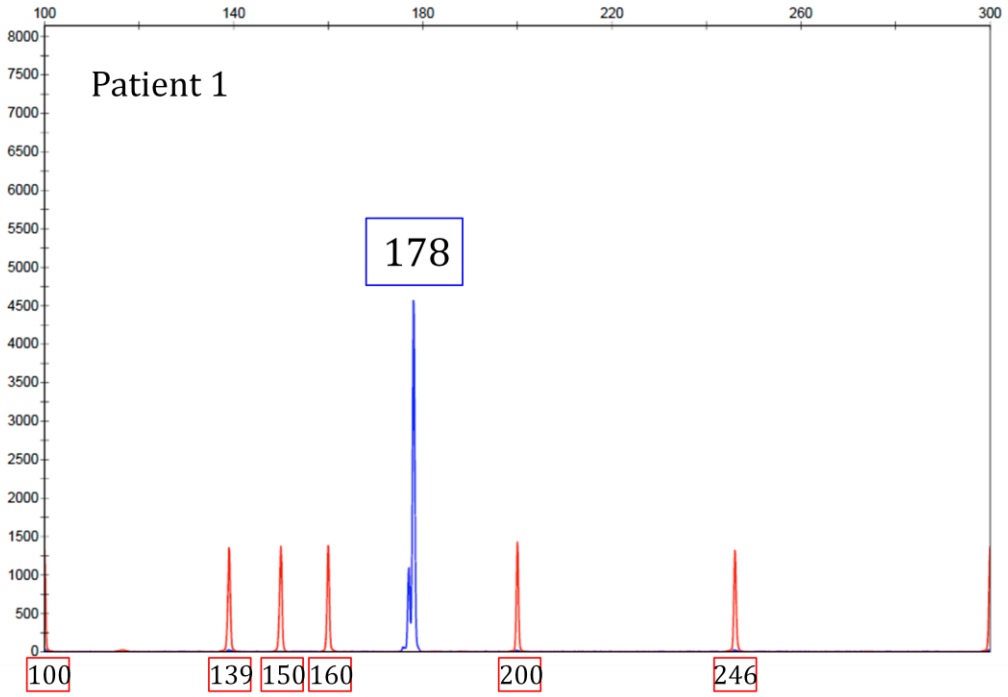
H


Negative control
100
140
180
220
260
300
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
100
139
150
160
200
246

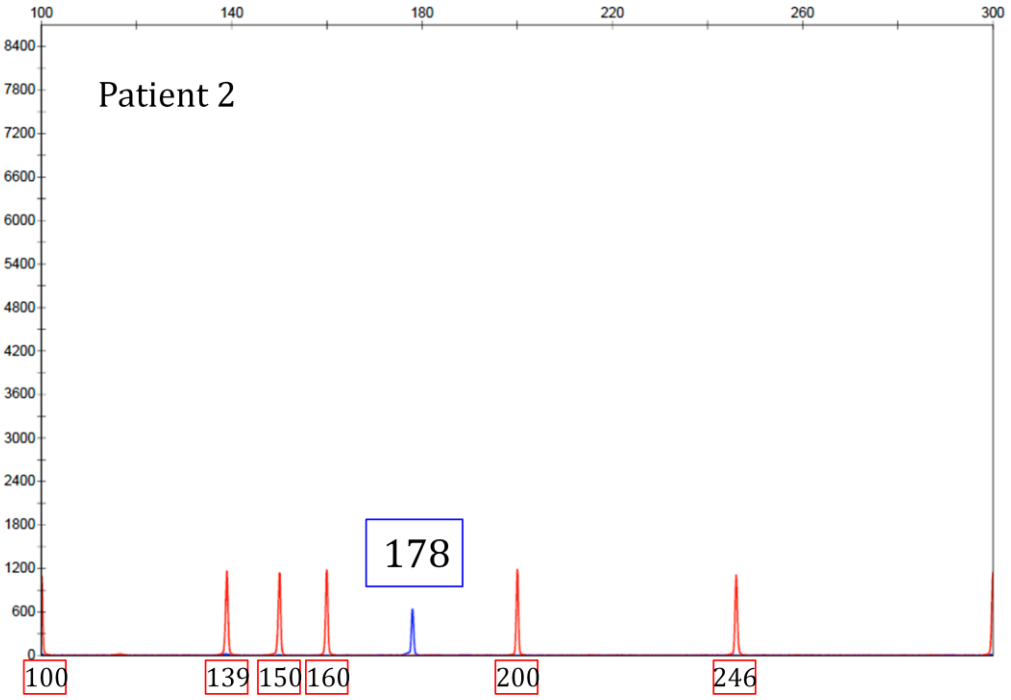
I


Fibroblast control
178
100
140
180
220
260
300
7000
6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
100
139
150
160
200
246

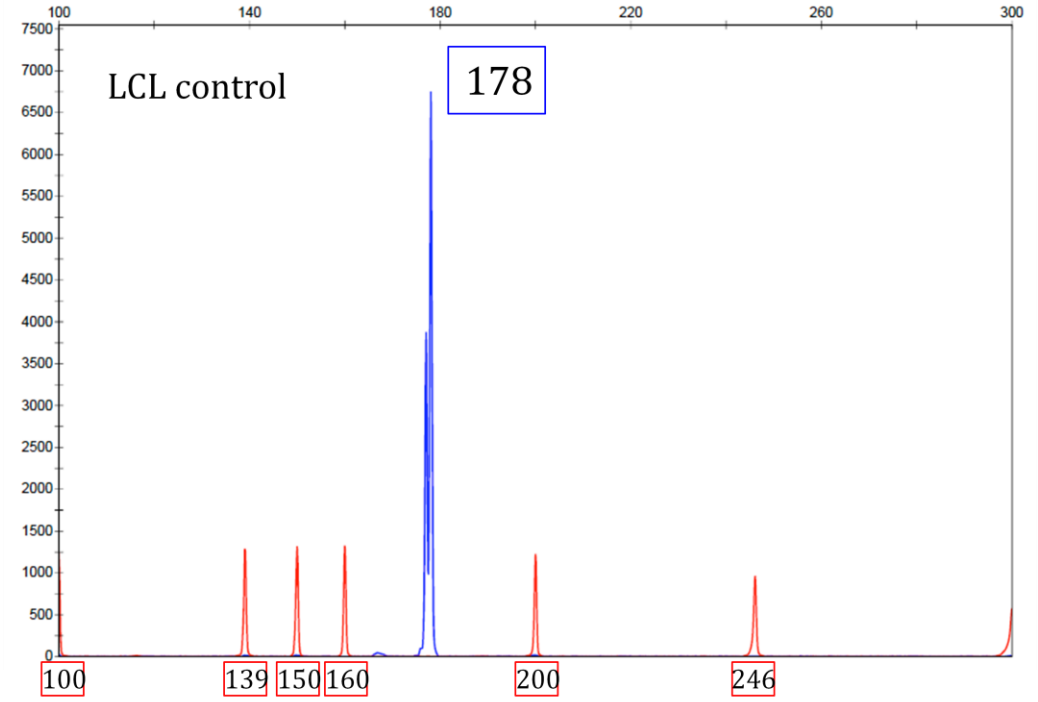
J


Patient 1
178
100
139
150
160
200
246

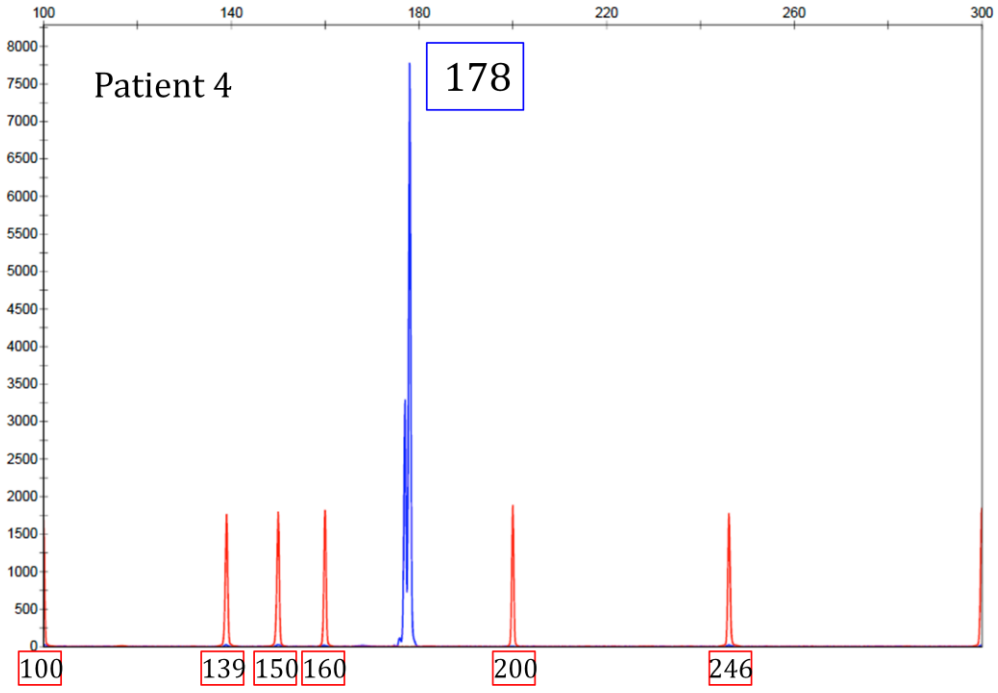
K


Patient 2
178
100
139
150
160
200
246

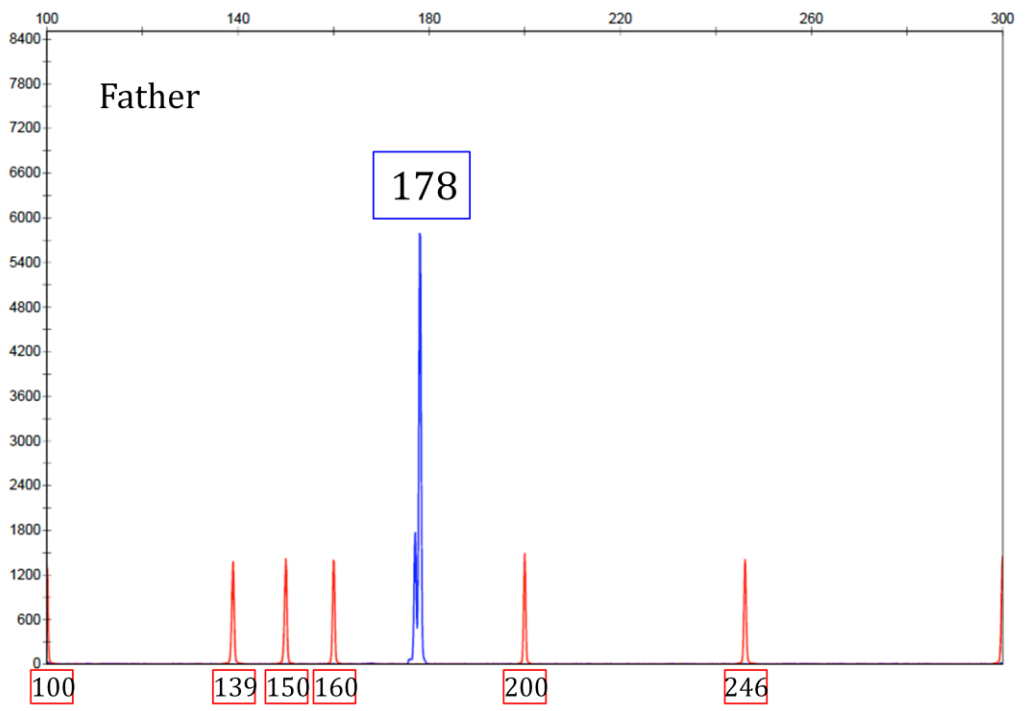
L


LCL control
178
100
140
180
220
260
300
100
139
150
160
200
246

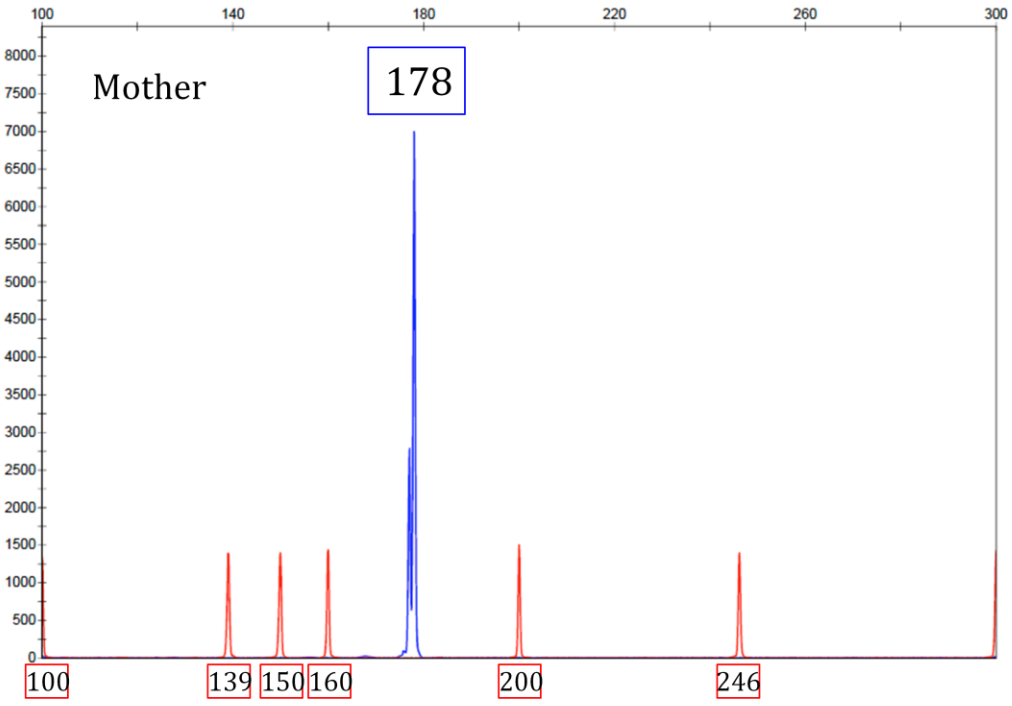
M

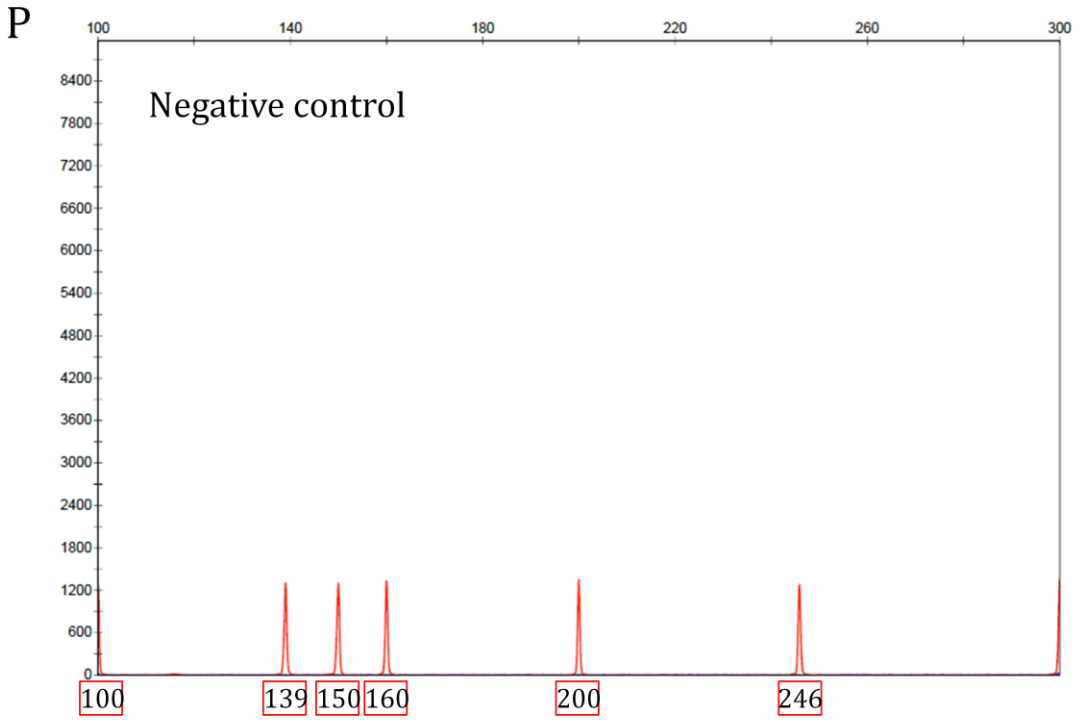

Patient 4
178
100
140
180
220
260
300
100
139
150
160
200
246

N

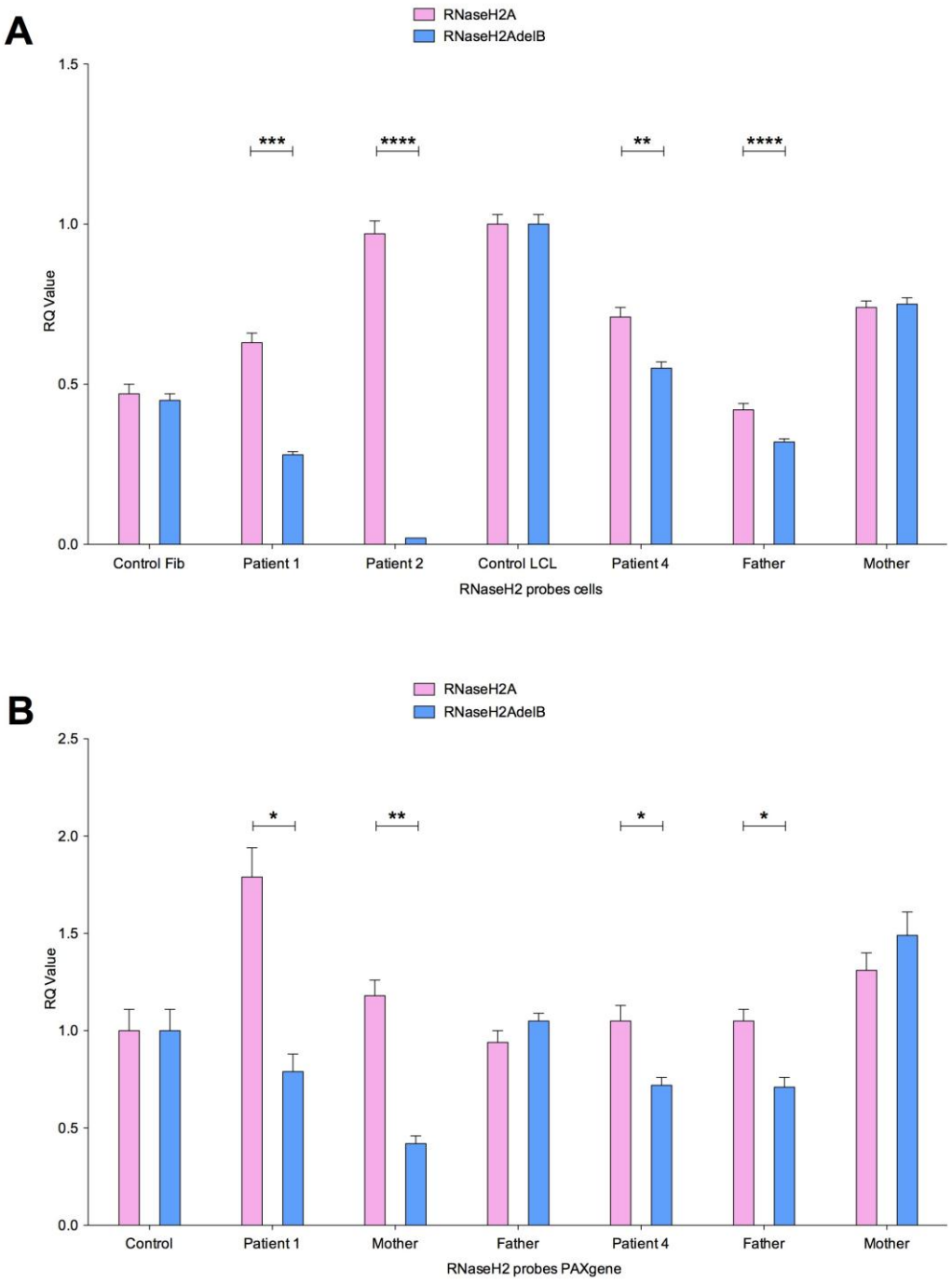

Father
178
100
139
150
160
200
246

O


Mother
178
100
139
150
160
200
246

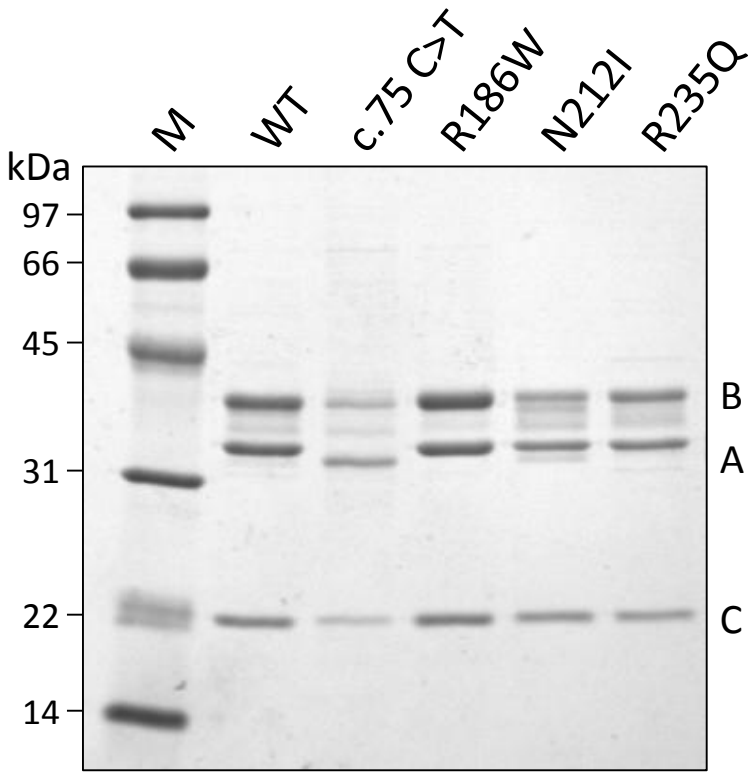


Supp. Figure S3. Chromatograms showing fluorescent RT-PCR products separated by denaturing capillary electrophoresis. **A-H:** RT-PCR products generated with primer pair 1 (RNaseH2A1F 5'FAM and RNaseH2AdelR2) and run at 1:7.5 dilution. Panel **C** shows product for patient 2 run at 1:7.5 dilution, and a zoomed in image of the product when run undiluted, showing an additional peak at 212bp apparently demonstrating the presence of a small amount of full-length product**. I-P:** RT-PCR products generated with primer pair 2 (RNaseH2AdelFnew 5'FAM and RNaseH2AdelRnew) and run at 1:20 dilution. Panels **H** and **P** show the no-template control RT-PCR reactions for each primer pair.

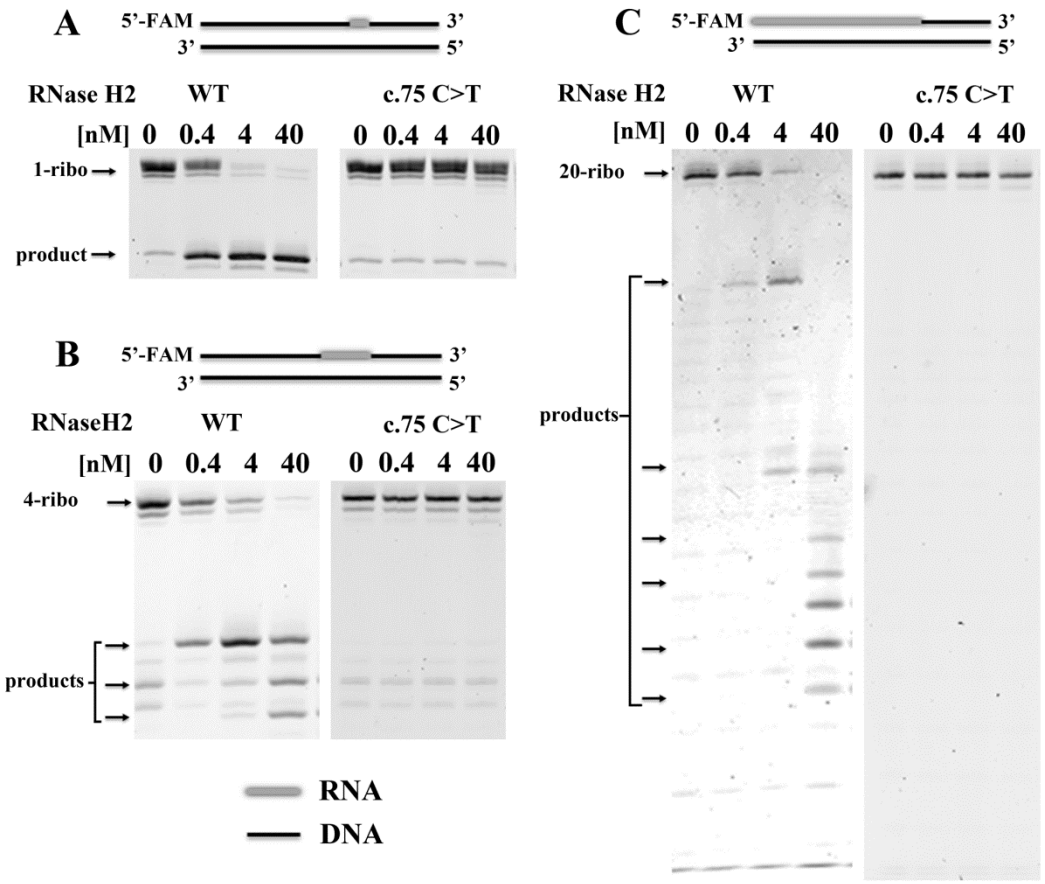


Supp. Figure S4. qPCR showing the relative abundance of RNASEH2A transcript with and without internal deletions in AGS patient cells and in whole blood. Taqman probes were used to determine the relative abundance of RNASEH2A (hs00197370-m1; detecting all transcripts at the exon 4/5 junction, including those with the deletion in exon 1) and RNASEH2A full-length transcript (custom probe RNaseH2AdelB only detects RNASEH2A transcripts without the deletion in exon 1). Values were normalised against HPRT1 and 18S, and expressed as RQ values relative to those found in control RNA. **A**:

RNA extracted from patient lymphoblastoid cells (patient 4 and parents) and fibroblasts (patients 1 and 2) compared to healthy control lymphoblastoid and fibroblast cell lines; **B**: RNA extracted from peripheral whole blood of patients and 5 controls. n=3; error bars indicate standard deviations. RQ is equal to $2^{-\Delta\Delta Ct}$, with $-\Delta\Delta Ct \pm$ standard deviations, that is the normalized fold change relative to a calibrator. **** $p<0.0001$; *** $p<0.001$; ** p 0.001 to 0.01; * p 0.01 to 0.05.



Supp. Figure S5. Truncated RNASEH2A protein, resulting from the c.75C>T mutation forms an intact recombinant RNase H2 complex. RNASEH2A (WT, with c.75C>T skipped region, and with Arg186Trp, Asn121Ile or Arg235Gln point mutations) was co-expressed with the B and C subunits in *E. coli*, and purified as described in Supplementary materials and methods. 4 µg of each purified RNase H2 complex was separated by SDS-PAGE and visualised by coomassie staining. Molecular weight markers (M) and the RNASEH2A, B and C subunits are annotated.



Supp. Figure S6. RNase H2 with RNASEH2A c.75C>T variant protein is catalytically inactive. 200 nM 1-ribo (**A**), 4-ribo (**B**) or 20-ribo (**C**) duplex polynucleotide substrates were incubated with increasing concentrations (0, 0.4 nM, 4 nM, 40 nM) WT or c.75C>T RNASEH2A as indicated. Reaction products were visualised by denaturing PAGE and quantified as described previously (Coffin, et al., 2011). Arrows show the positions of substrates and products.

Supp. Table S1. *RNASEH2A* variants identified in affected individuals and their parents

Patient	Number affected	*RNASEH2A* mutations		
		Affected	**Father**	**Mother**
1	2	c.75C>T (p.Arg25Arg) het + c.704G>A (p.Arg235Gln) het	c.704G>A (p.Arg235Gln) het	c.75C>T (p.Arg25Arg) het
2	2	c.75C>T (p.Arg25Arg) hom	c.75C>T (p.Arg25Arg) het	c.75C>T (p.Arg25Arg) het
3	1	c.69G>A (p.Val23Val) het + c.556C>T (p.Arg186Trp) het	c.556C>T (p.Arg186Trp) het	c.69G>A (p.Val23Val) het
4	1	c.69G>A (p.Val23Val) het + c.635A>T (p.Asn212Ile) het	c.69G>A (p.Val23Val) het	c.635A>T (p.Asn212Ile) het

Supp. Table S2. Primers used for amplification and sequencing of *RNASEH2A* cDNA

Primer name	Primer sequence
RNASEH2A1F	GCTCCTGCAGTATTAGTTCTTG
RNASEH2A7R	TACGTGTGGTTCTCCTTAAACA
cDNA1R	GAGGCAGGGGACAATAACAG
cDNA2F	AATGGAGGACACGGACTTTG
cDNA2R	ATGTCTCTGGCATCCCTACG
cDNA3F	TCAGGAGGGACTCAGGAAGA

Supp. Table S3. Primers used for RT-PCR of *RNASEH2A*

Primer name	Primer sequence	Product size cDNA	Product size if deletion	Product size genomic DNA
RNASEH2A 1F*	GCTCCTGCAGTATTAGTTCTTG	214bp	153bp (c.69G>A)	421bp
RNASEH2A delR2	GAGGCAGGGGACAATAACAG		160bp (c.75C>T)	
RNASEH2A delFnew*	GTCCTGGGCGTCGATGAG	178bp	No amplification	511bp
RNASEH2A delRnew	CAAAGTCCGTGTCCTCCATT			

*For fluorescent RT-PCR, primers were purchased with a 5'FAM label

Supp. Table S4. Peak size and area of fluorescent RT-PCR products, primer set 1

Primer set 1. 1:7.5 dilution		
Patient	**Peak size**	**Peak area**
Fibroblast control	212	2888
Patient 1	160	7720
	212	1517
Patient 2	160	9627
Patient 2*	160	10593
	212	621
LCL control	212	8639
Patient 4	153	2781
	212	7010
Father	153	2088
	212	3964
Mother	212	6413

*when product run undiluted

Supp. Table S5. Peak size and area of fluorescent RT-PCR products, primer set 2

Primer set 2 1:20 dilution		
Patient	**Peak size**	**Peak area**
Fibroblast control	178	3222
Patient 1	178	4574
Patient 2	178	646
LCL control	178	6752
Patient 4	178	7782
Father	178	5795
Mother	178	7002

Supp. Materials and Methods

Subjects

Subjects with a clinical diagnosis of AGS from four pedigrees were recruited internationally through collaborating physicians. Clinical, neuroradiological and laboratory data were provided by the clinician responsible for the care of the patient. With consent, blood samples were obtained from affected children, their parents, and unaffected siblings. The study was approved by the Leeds (East) Multi-centre Research Ethics Committee (reference number 07/Q1206/7).

Mutation analysis

Genomic DNA was extracted from peripheral-blood leukocytes by standard methods. PCR amplification of all coding exons and conserved splice sites of *TREX1*, *RNASEH2A*, *RNASEH2B*, *RNASEH2C* and *SAMHD1* was performed as previously described (Rice, et al., 2007; Rice, et al., 2009). Purified PCR-amplification products were sequenced using dye-terminator chemistry as reported previously (Rice, et al., 2007). One hundred European control alleles were sequenced for all five genes and, in the case of *RNASEH2C*, 100 South Asian alleles were also analyzed.

RNA extraction and reverse transcription PCR

RNA was extracted from subject fibroblast and lymphoblastoid cell pellets using the RNeasy mini kit (Qiagen). Total RNA was extracted from whole blood using a PAXgene (PreAnalytix) RNA isolation kit. RNA was treated using the Ambion DNaseI DNA-free kit, and concentration assessed using a spectrophotometer (FLUOstar Omega, Labtech).

Reverse-transcription of 2 μg of RNA was carried out using the QIAGEN OneStep RT-PCR kit, with primers RNASEH2A1F and RNASEH2A7R (30 minutes at 50^{o}C followed by 40 cycles with an annealing temperature of 60^{o}C and 1 minute elongation). The PCR product (1019 bp for full-length) was sequenced using cDNA sequencing primers (Supp Table S2). Additional RT-PCR amplification was carried out using primers RNaseH2A1F and RNaseH2AdelR2 (primers situated either side of deletion (Supp. Table

S3 and Supp. Figure S2) (32 cycles of amplification)); and primers RNaseH2AdelFnew and RNaseH2AdelRnew (forward primer situated within deleted region, reverse primer in exon 3 (Supp. Table S3 and Supp. Figure S2) (32 cycles of amplification)). RT-PCR amplification using 5'FAM labeled forward primers was carried out for 30 cycles, and products separated by denaturing capillary electrophoresis (ABI Prism 3130xl Genetic Analyser). Data was analysed using GeneMapper 4.1 Software (Applied Biosystems).

qPCR

Quantitative PCR (qPCR) analysis was performed using the TaqMan Universal PCR Master Mix (Applied Biosystems) and cDNA derived from 40 ng total RNA. The levels of target transcripts were measured using a Taqman probe covering the exon 4/5 junction of *RNASEH2A* (Hs00197370_m1) and a custom probe designed against the region of *RNASEH2A* deleted in cells carrying the c.69G>A (p.Val23Val) and c.75C>T (p.Arg25Arg) mutations (RNaseH2AdelB: RNASEH2A_F CATCTGTTATTGTCCCCTGCCT, RNASEH2A_R CCCGCTCGCTCTCCAATAG and RNASEH2A_M FAM TTGAGTCTGCCACTTTC). Relative abundance was calculated after normalization against *HPRT1* (Hs03929096_g1) and 18s (Hs999999001_s1) using the Applied Biosystems StepOne Software v2.1, and Applied Biosystems Data Assist Software v.3.01.

Cell culture

Human skin fibroblasts were cultured at 37^{o}C in 5% CO_2 in D-MEM (Dulbecco's Modified Eagle medium) supplemented with GlutaMAX-I, 4500 mg/L D-Glucose, sodium pyruvate, 100 IU/ml Penicillin, 100 ug/ml streptomycin (all from Invitrogen) and 10% Foetal Bovine Serum Gold (PAA Laboratories Ltd). Human lymphoblastoid cell lines for family 4 were cultured at 37^{o}C in 5% CO_2 in RPMI-1640 medium supplemented with GlutaMAX-I, 100 IU/ml Penicillin, 100 ug/ml streptomycin (all from Invitrogen) and 10% Foetal Bovine Serum Gold (PAA Laboratories Ltd).

Recombinant RNase H2 expression, purification and enzyme activity assays

The human RNase H2 and mutant heterotrimer complexes were expressed and purified as previously described (Coffin, et al., 2011; Perrino, et al., 2009). Point mutations c.556C>T (p.Arg186Trp) and c.704G>A (p.Arg235Gln) were introduced into the pET-A expression plasmid containing the *RNASEH2A* gene using QuikChange site directed mutagenesis (Kunkel, 1985). The *RNASEH2A* coding sequence with either the c.635A>T (p.Asn212Ile) or c.75C>T (p.Arg25Arg) mutation was cloned into the pET-A expression plasmid after PCR amplification using template cDNA from peripheral blood leukocytes of AGS patients. Purified enzymes were concentrated and stored at -80°C in 20 mM Tris-HCl (pH 7.5), 50 mM NaCl, and 10% glycerol. Enzyme concentrations were calculated using the molar extinction coefficient (83,030 $M^{-1}cm^{-1}$) at A_{280}.

Relative RNase H2 activities were measured in reactions (30 µl) containing 20 mM Tris–HCl pH 7.5, 5 mM $MgCl_2$, 2 mM dithiothreitol, 50 µg/ml bovine serum albumin, 200 nM of duplex RNA/DNA hybrids containing 1, 4, or 20 ribonucleotides and increasing amounts of RNase H2 (0, 0.4, 4, 40, 400 nM) as previously described (Coffin, et al., 2011). Reactions were incubated at 25°C for 20 min, quenched with three volumes of 100% ethanol, and dried *in vacuo*. Pellets were resuspended in 9 µl of formamide and heated at 95°C for 5 min. Reaction products from a sample (5 µl) were separated on 23% urea-polyacrylamide gels and visualized using a Storm PhosphorImager (GE Healthcare).

Western analysis

Cells were lysed as previously described (Crow, et al., 2006), and proteins were separated by SDS-PAGE on 4-12% NuPAGE gels (Invitrogen). After Western blotting, RNASEH2B and RNASEH2C were detected with sheep polyclonal antibody raised against purified recombinant RNase H2 complex (Reijns, et al., 2012), RNASEH2A with rabbit polyclonal antibody (Origene), and GAPDH with HRP-conjugated rabbit anti-GAPDH (Abcam).

Cellular RNase H activity assay

To determine levels of cellular RNase H2 activity, assays were performed using cell lysates and a fluorometric assay as previously described (Reijns, et al., 2011). In each

experiment, measurements were performed in triplicate for each cell lysate. Experiments were performed in duplicate, using independent cell lysates from cells harvested on different days. Briefly, 10 µM of 3' fluorescein-labeled oligonucleotides (GATCTGAGCCTGGGaGCT or gatctgagcctgggagct; uppercase DNA, lowercase RNA) were annealed to a complementary 5' DABCYL-labeled DNA oligonucleotide (Eurogentec) in 60 mM KCl, 50 mM Tris-HCl pH 8, by heating for 5 min at 95°C followed by slow cooling to room temperature. Total protein cell lysates were prepared as for Western analysis and concentrations determined by Bradford assay (Bio-Rad). Reactions were performed in 100 µl of buffer (60 mM KCl, 50 mM Tris-HCl pH 8, 10 mM $MgCl_2$, 0.01% BSA, 0.01% Triton X-100) with 250 nM substrate and 50 ng/µl of total protein in 96-well flat-bottomed plates at 24 ± 2°C for 45-90 min. Fluorescence was read for 100 ms using a VICTOR2 1420 multilabel counter (Perkin Elmer), with a 480 nm excitation filter and a 535 nm emission filter.

Supp. References

Coffin SR, Hollis T, Perrino FW. 2011. Functional Consequences of the RNase H2A Subunit Mutations That Cause Aicardi-Goutieres Syndrome. J Biol Chem 286:16984-91.

Crow YJ, Leitch A, Hayward BE, Garner A, Parmar R, Griffith E, Ali M, Semple C, Aicardi J, Babul-Hirji R and others. 2006. Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutieres syndrome and mimic congenital viral brain infection. Nat Genet 38:910-6.

Kunkel TA. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc Natl Acad Sci U S A 82:488-92.

Perrino FW, Harvey S, Shaban NM, Hollis T. 2009. RNaseH2 mutants that cause Aicardi-Goutieres syndrome are active nucleases. J Mol Med (Berl) 87:25-30.

Reijns MA, Bubeck D, Gibson LC, Graham SC, Baillie GS, Jones EY, Jackson AP. 2011. The Structure of the Human RNase H2 Complex Defines Key Interaction Interfaces Relevant to Enzyme Function and Human Disease. J Biol Chem 286:10530-9.

Reijns MA, Rabe B, Rigby RE, Mill P, Astell KR, Lettice LA, Boyle S, Leitch A, Keighren M, Kilanowski F and others. 2012. Enzymatic removal of ribonucleotides from DNA is essential for Mammalian genome integrity and development. Cell 149:1008-22.

Rice G, Patrick T, Parmar R, Taylor CF, Aeby A, Aicardi J, Artuch R, Montalto SA, Bacino CA, Barroso B and others. 2007. Clinical and molecular phenotype of Aicardi-Goutieres syndrome. Am J Hum Genet 81:713-25.

Rice GI, Bond J, Asipu A, Brunette RL, Manfield IW, Carr IM, Fuller JC, Jackson RM, Lamb T, Briggs TA and others. 2009. Mutations involved in Aicardi-Goutieres syndrome implicate SAMHD1 as regulator of the innate immune response. Nat Genet 41:829-32.

Rychlik MP, Chon H, Cerritelli SM, Klimek P, Crouch RJ, Nowotny M. 2010. Crystal structures of RNase H2 in complex with nucleic acid reveal the mechanism of RNA-DNA junction recognition and cleavage. Mol Cell 40:658-70.