

Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutières syndrome and mimic congenital viral brain infection

Yanick J Crow^{1,3}, Andrea Leitch², Bruce E Hayward¹, Anna Garner², Rekha Parmar¹, Elen Griffith², Manir Ali¹, Colin Semple², Jean Aicardi⁴, Riyana Babul-Hirji⁵, Clarisse Baumann⁶, Peter Baxter⁷, Enrico Bertini⁸, Kate E Chandler⁹, David Chitayat⁵, Daniel Cau¹⁰, Catherine Déry¹¹, Elisa Fazzi¹², Cyril Goizet¹³, Mary D King¹⁴, Joerg Klepper¹⁵, Didier Lacombe¹³, Giovanni Lanzi¹², Hermione Lyall¹⁶, María Luisa Martínez-Frías¹⁷, Michèle Mathieu¹⁸, Carole McKeown¹⁹, Anne Monier⁴, Yvette Oade²⁰, Oliver W Quarrell²¹, Christopher D Rittey²², R Curtis Rogers²³, Amparo Sanchis²⁴, John B P Stephenson²⁵, Uta Tacke²⁶, Marianne Till²⁷, John L Tolmie²⁸, Pam Tomlin²⁹, Thomas Voit¹⁵, Bernhard Weschke³⁰, C Geoffrey Woods³¹, Pierre Lebon³², David T Bonthron¹, Chris P Ponting³³ & Andrew P Jackson²

Aicardi-Goutières syndrome (AGS) is an autosomal recessive neurological disorder, the clinical and immunological features of which parallel those of congenital viral infection. Here we define the composition of the human ribonuclease H2 enzyme complex and show that AGS can result from mutations in the genes encoding any one of its three subunits. Our findings demonstrate a role for ribonuclease H in human neurological disease and suggest an unanticipated relationship between ribonuclease H2 and the antiviral immune response that warrants further investigation.

AGS (MIM 225750) is an autosomal recessive genetic disorder that is phenotypically similar to *in utero* viral infection¹. Severe neurological dysfunction becomes clinically apparent in infancy, and manifests as progressive microcephaly, spasticity, dystonic posturing, profound psychomotor retardation and often death in early childhood^{2,3}. Intracranial calcification, white matter destruction and brain atrophy

are evident on neuroimaging (**Supplementary Fig. 1** online), similar to the radiological findings seen in congenital cytomegalovirus⁴ and rubella syndromes⁵ and transplacentally acquired HIV encephalopathy^{6,7}. Outside the nervous system, thrombocytopenia, hepatosplenomegaly and elevated hepatic transaminases along with intermittent fever may also erroneously suggest an infective process³.

¹Leeds Institute of Molecular Medicine, University of Leeds, St. James's University Hospital, Leeds, LS9 7TF, UK. ²UK Medical Research Council (MRC) Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh, EH4 2XU, UK. ³Department of Paediatrics, St. Luke's Hospital, Little Horton Lane, Bradford BD5 0NA, UK. ⁴Service de Neurologie Pédiatrique et des Maladies Métaboliques, Hôpital Robert Debré, 48 Boulevard Sérurier, 75935 Paris Cedex 19, France. ⁵Division of Clinical and Metabolic Genetics, The Hospital for Sick Children, 555 University Avenue, Toronto, ON M5G 1X8, Canada. ⁶Unité de Génétique Clinique, Hôpital Robert Debré, Boulevard Sérurier, 75019 Paris, France. ⁷Sheffield Children's Hospital, Western Bank, Sheffield, S10 2TH, UK. ⁸Unit of Molecular Medicine, Bambino Gesù Children's Hospital, P.zza S. Onofrio 4, 00165 Rome, Italy. ⁹University Department of Medical Genetics and Regional Genetic Service, St. Mary's Hospital, Central Manchester Healthcare Trust, Hathersage Road Manchester. M13 0JH, UK. ¹⁰Service de Pédiatrie/Néonatalogie, Hôpital William Morey BP 120, 71120 Chalon sur Saone, France. ¹¹Centre de Santé et de Services Sociaux de Beauce, 1515, 17ième rue, St-Georges, Québec G5Y 4T8, Canada. ¹²Department of Child Neurology and Psychiatry, IRCCS C. Mondino Institute of Neurology, University of Pavia, Pavia, Italy. ¹³Medical Genetics Department, Pellegrin Teaching Hospital, 33076 Bordeaux, France. ¹⁴Department of Paediatric Neurology, The Children's University Hospital, Temple Street, Dublin 1, Ireland. ¹⁵University of Essen, Department of Pediatric Neurology, Hufelandstrasse 55, D-45122, Essen, Germany. ¹⁶Department of Paediatric Infectious Diseases, St. Mary's Hospital, Praed Street, London W2 1NY, UK. ¹⁷Centro de Investigación sobre Anomalías Congénitas, Instituto de Salud Carlos III, Sinesio Delgado 6, Pabellón 6, Madrid 28029, Spain. ¹⁸Unité de Génétique Clinique, Hôpital Nord, Place Victor Pauchet, 80054 Amiens Cedex, France. ¹⁹Clinical Genetics Unit, Norton Court, Birmingham Women's Hospital, Birmingham, B15 2TG, UK. ²⁰Department of Paediatrics, Calderdale Royal Hospital, Salterhebble, Halifax, HX3 0PW, UK. ²¹Department of Clinical Genetics, Sheffield Children's Hospital, Sheffield, S10 2TH, UK. ²²Ryegate Children's Centre, Tapton Crescent Road, Sheffield S10 5DD, UK. ²³Greenwood Genetic Center, 1 Gregor Mendel Circle, Greenwood, South Carolina 29646, USA. ²⁴Servicio de Pediatría, Hospital Universitario Doctor Peset, Avda Gaspar Aguilar 90, 46017 Valencia, Spain. ²⁵Fraser of Allander Neurosciences Unit, Royal Hospital for Sick Children, Yorkhill, Glasgow, G3 8SJ, UK. ²⁶University Children's Hospital, Mathildenstr. 1, 79106 Freiburg, Germany. ²⁷Department of Genetics, Debrousse Hospital, 69322 Lyon cedex 05, France. ²⁸Duncan Guthrie Institute of Medical Genetics, Royal Hospital for Sick Children, Yorkhill, Glasgow, G3 8SJ, UK. ²⁹Department of Paediatric Neurology, Royal Preston Hospital, Sharoe Green Lane, Fulwood, Preston PR2 9HT, UK. ³⁰Department of Neuropediatrics, Charité Medical School, Augustenburger Platz, D-13353 Berlin, Germany. ³¹East Anglian Medical Genetics Service, Addenbrooke's NHS Trust Box 134, Hills Road Cambridge CB2 2QQ, UK. ³²Service de Virologie, Hôpital Cochin, St. Vincent de Paul, 82 Avenue Denfert Rochereau, 75674, Paris and Université René Descartes, Paris, France. ³³MRC Functional Genetics Unit, Department of Physiology, Anatomy & Genetics, University of Oxford, South Parks Road, Oxford OX1 3QX, UK. Correspondence should be addressed to Y.J.C. (yanickcrow@mac.com) and A.P.J. (andrew.jackson@hgu.mrc.ac.uk).

Received 19 April; accepted 12 June; published online 16 July 2006; doi:10.1038/ng1842

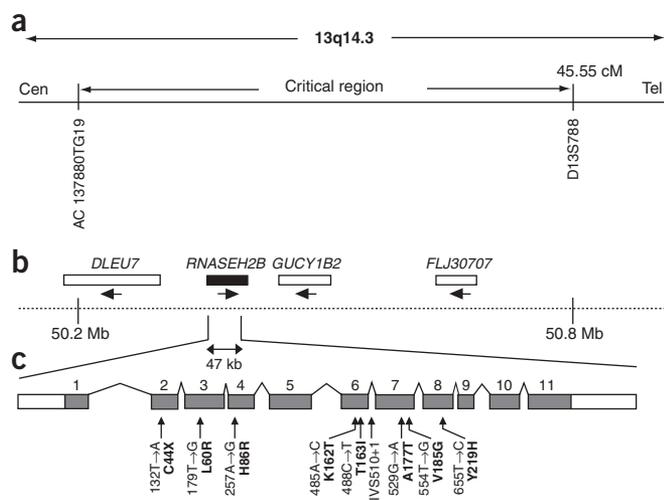


Figure 1 Schematic of *AGS2* critical region and *RNASEH2B* gene depicting location of identified mutations. **(a)** Genetic map of chromosome 13q14.1 depicting the refined *AGS2* locus, defined by overlapping homozygous chromosomal segments in non-consanguineous families, and extending between microsatellite markers AC137880TG19 and D13S788. **(b)** The physical map of the 571-kb critical interval contains four annotated genes (UCSC Genome Browser May 2004 assembly). **(c)** *RNASEH2B* spans 47 kb of genomic sequence in 11 exons and encodes a 308-amino acid protein. Coding sequence shaded gray. Locations of mutations are indicated by vertical and slanted arrows, with mutation position enumerated relative to the translational start site and the corresponding amino acid change in bold.

Significant production of interferon α in the cerebrospinal fluid is a feature of AGS and many congenital viral infections^{3,8–11}. Increased numbers of lymphocytes are also present in the cerebrospinal fluid², and perivascular lymphocytic infiltrates have been reported in small cerebral vessels and meninges^{3,12}. Hence, given its clinical and immunological features, it is not surprising that AGS can be mistaken for congenital infection and, conversely, that exclusion of perinatal infection is a diagnostic criterion for AGS.

Previously, we have identified two AGS loci, *AGS1* on chromosome 3p21 (refs. 13,14) and *AGS2* on chromosome 13q14.3 (ref. 15), and we inferred that additional loci remained to be identified. Here we report the refinement of the *AGS2* locus and define a third locus (*AGS3*) on 11q13.2. Using a combination of positional cloning, homology prediction and functional studies, we then determined the human orthologs for each of the three *Saccharomyces cerevisiae* ribonuclease (RNase) H2 enzyme subunits and identified pathogenic mutations in the genes encoding all three: *AGS2* (*RNASEH2B*, FLJ11712), *AGS3* (*RNASEH2C*, AYP1) and *AGS4* (*RNASEH2A*).

RESULTS

AGS2 locus refinement

We performed high-density genotyping of microsatellite markers in a panel of ten families to refine the *AGS2* locus. We found that two non-consanguineous families (F8 and F10; **Supplementary Fig. 2** online) had small overlapping regions of homozygous markers. Such regions are likely to represent autozygous chromosomal segments^{16,17}, which in rare autosomal recessive disorders have a high probability of containing the disease gene¹⁸. On the basis of this overlap, we refined the *AGS2* critical interval to a 571-kb region on chromosome 13q14.3 (**Fig. 1**). We then sequenced the coding exons of all four annotated genes in this interval.

RNASEH2B is the *AGS2* gene

We identified missense sequence changes in *RNASEH2B* in seven of the families screened but did not observe any pathogenic mutations in *DLEU7*, *GUCY1B2* or *FLJ30707*. Subsequent mutation screening of a larger cohort of affected individuals identified a total of 18 families with mutations in *RNASEH2B* (**Table 1** and **Fig. 1**). Most mutations were missense, with two (the mutations resulting in A177T and V185G) found recurrently in groups of different national origin. All missense mutations resulted in nonconservative replacement of residues that are conserved among mammals (**Supplementary Fig. 3** online), with the exception of a single conservative residue change¹⁹ (Y219H). However this residue, Tyr219, is conserved as far back as slime mold (*Dictyostelium*

Table 1 *RNASEH2B* mutations identified in 18 Aicardi-Goutières syndrome families

Family	Ancestry	Nucleotide alterations	Amino acid alterations	Exon	Segregation	Parental consanguinity
F1	Moroccan	485A→C	K162T	6	Hom, M, P	Yes
F3	Italian	554T→G	V185G	7	Hom, M, P	Yes
F4	Algerian	529G→A	A177T	7	Hom, M, P	Yes
F5	Moroccan	529G→A	A177T	7	Hom, M, P	Yes
F8	Italian	554T→G	V185G	7	Hom, M, P	No
F9	Irish	529G→A	A177T	7	Hom, M, P	Yes
F10	Moroccan	529G→A	A177T	7	Hom, nps	No
F11	Mixed Tunisian/Algerian	[257A→G]+[529G→A]	[H86R]+[A177T]	4; 7	Het, M, P	No
F12	Italian	[488C→T]+[529G→A]	[T163I]+[A177T]	6; 7	Het, P	No
F13	Italian	529G→A	A177T	7	Hom, M, P	No
F14	French Canadian	529G→A	A177T	7	Hom, M, P	No
F15	British	[IVS510+1G→A]+[529G→A]	[Exon 6 splice donor site]+[A177T]	Intron 6; 7	Het, M, P	No
F16	German	529G→A	A177T	7	Hom, nps	No
F17	American, European origin	[132T→A]+[655T→C]	[C44X]+[Y219H]	2; 8	Het, M, P	No
F18	American, European origin	[179T→G]+[529G→A]	[L60R]+[A177T]	3; 7	Het, M, P	No
F19	Italian	[488C→T]+[529G→A]	[T163I]+[A177T]	6; 7	Het, nps	No
F20	Italian	[488C→T]+[529G→A]	[T163I]+[A177T]	6; 7	Het, M, P	No
F21	Mixed European Canadian and Hungarian	[529G→A]+[554T→G]	[A177T]+[V185G]	7	Het, M, P	No

Nucleotides are numbered from the A of the initiation codon (ATG) in the nucleotide sequence NM_024570. Abbreviations: fs, frameshift. Hom, homozygous in affected individual; Het, heterozygous in affected individual; M, mutation identified in mother; P, mutation identified in father; nps, no parental samples.

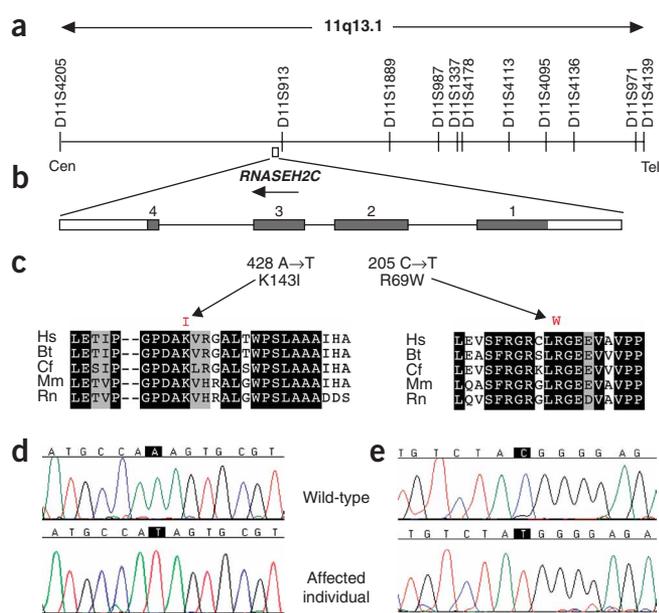


Figure 2 Schematic of *AGS3* region, the *RNASEH2C* gene, its mutations and sequence conservation in other species. **(a)** Genetic map of chromosome 11q13.1. The critical region is defined by linkage disequilibrium data in the Pakistani families and lies between D11S4205 and D11S987. The *RNASEH2C* gene lies at 65.2 Mb spanning 1.4 kb of genomic sequence in a telomeric-to-centromeric orientation on the minus strand. **(b)** The *RNASEH2C* gene structure comprises four exons (coding sequence in gray) and encodes a 164-amino acid protein. The location of mutations are indicated by vertical lines, with mutation position enumerated relative to the translational start site, and the corresponding amino acid change in bold. **(c)** Both mutations affect residues conserved in mammals. Sequence alignments of mammalian *RNASEH2C* proteins that immediately flank the substituted residues. Hs, *Homo sapiens*; Bt, *Bos Taurus*; Cf, *Canis familiaris*; Mm, *Mus musculus*; Rn, *Rattus norvegicus*. Substituted amino acids are shown in red above sequence alignment. **(d,e)** Sequence electropherograms of *RNASEH2C* mutations 428A→T **(d)** and 205C→T **(e)**.

the possibility that additional RNase H components might also be mutated in AGS.

RNase H2

RNase H enzymes endonucleolytically cleave ribonucleotides from RNA:DNA duplexes. Two classes of RNase H exist, type 1 (or I) and 2 (or II), each possessing distinctive biochemical properties²². RNase H2 is the major source of RNase H activity in mammalian cells and yeast^{23–25}. It is proposed to function in the removal of lagging-strand Okazaki fragment RNA primers during DNA replication²⁶ and in the excision of single ribonucleotides from DNA:DNA duplexes²⁵. In *S. cerevisiae*, Rnh2Bp and Rnh2Cp copurify with the catalytic subunit Rnh2Ap and together are sufficient to reconstitute RNase H2 activity²¹. In humans, the RNase H2 catalytic subunit, RNASEH2A, has been identified by biochemical purification and shows significant sequence similarity with its yeast ortholog^{23,27,28}. However, no orthologs outside yeast had been identified previously for the Rnh2Bp and Rnh2Cp subunits²¹. Conversely, no *S. cerevisiae* ortholog of the human protein RNASEH2C (AYP1, AF312034), found to copurify with the RNASEH2A protein²³, has been recognized.

RNASEH2C is the *AGS3* gene

We performed genome-wide homozygosity mapping on six consanguineous families (five Pakistani (F30–34) and one Bangladeshi (F35)) using SNP microarrays. Together with subsequent microsatellite genotyping (**Supplementary Fig. 4** online), this identified a new locus, *AGS3*, on chromosome 11q13.2, with a maximum multipoint LOD score of 4.54 at 66.8 cM (Marshfield genetic map). Furthermore, genotypes in the five Pakistani families suggested the presence of an ancestral haplotype (**Supplementary Fig. 4**), which, along with the small homozygous region in family F30, refined the *AGS3* critical region to a 4.9-cM interval between D11S4205 and D11S987.

discoideum). We identified nonsense mutations in two families: a stop codon in exon 2 (F17), and a splice donor site mutation in intron 6 (F15). In both these cases, affected individuals were compound heterozygotes, with the second mutation being a missense change. The observed mutation spectrum therefore suggests that the mutational consequences will be hypomorphic rather than a complete loss of RNASEH2B protein function. Mutations segregated with the disease in all families, and all available parents were heterozygous for the mutations. At least 160 European control alleles were genotyped for each mutation. Only for the most common amino acid substitution, A177T, did we find one heterozygous individual (out of 241 samples tested), and this individual is thus likely to be a carrier for AGS.

RNASEH2B is the ortholog of yeast Rnh2Bp

RNASEH2B encodes a 308-amino acid protein of previously undefined function. Semi-quantitative RT-PCR indicates that *RNASEH2B* is detectable in a wide range of human tissues, suggesting ubiquitous expression (data not shown).

A database search using PSI-BLAST²⁰ showed significant similarity ($E = 4 \times 10^{-5}$) between human *RNASEH2B* and an *S. cerevisiae* protein, Rnh2Bp (Rnh202), after four search iterations (**Supplementary Fig. 3**). Rnh2Bp is an essential component of the yeast RNase H2 enzyme complex²¹. This observation indicated that *RNASEH2B* might function in an equivalent complex in mammals and raised

Table 2 *RNASEH2C* mutations identified in six Aicardi-Goutières syndrome families

Family	Ancestry	Nucleotide alterations	Amino acid alterations	Exon	Segregation	Parental consanguinity
F30	Pakistani	205C→T	R69W	2	Hom, M, P	Yes
F31	Pakistani	205C→T	R69W	2	Hom, M, P	Yes
F32	Pakistani	205C→T	R69W	2	Hom, M, P	Yes
F33	Pakistani	205C→T	R69W	2	Hom, M, P	Yes
F34	Pakistani	205C→T	R69W	2	Hom, M, P	Yes
F35	Bangladeshi	428A→T	K143I	3	Hom, M, P	Yes

Nucleotides are numbered from the start of the initiation codon (ATG) (for the protein transcript AF312034). Abbreviations: Hom, homozygous in affected individual; M, mutation identified in mother; P, mutation identified in father.

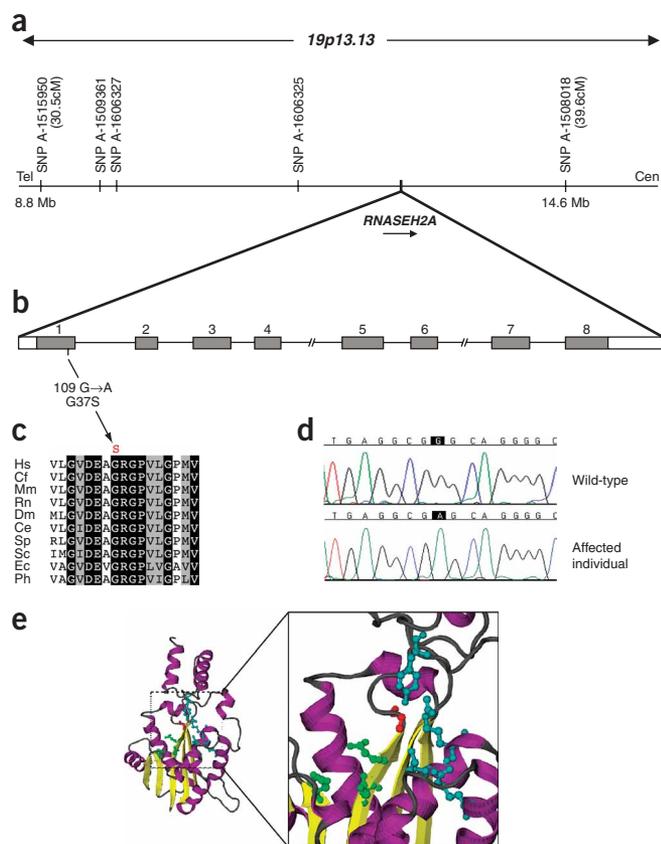


Figure 3 The *RNASEH2A* gene, genomic location, gene structure and mutation location. **(a)** Genetic map of chromosome 19p13.13. *RNASEH2A* lies in a region where two affected children of consanguineous second-cousin parents (family F39) share homozygous SNPs (SNP A-1509361, 1606327, 1606325), defining a region of potential genetic linkage between SNP A-1515950 and A-1508018. **(b)** *RNASEH2A* gene structure. *RNASEH2A* lies at 12.8 Mb (UCSC Genome Browser May 2004 assembly), spanning 7 kb of genomic sequence. It comprises eight exons and encodes a 299-amino acid protein. **(c)** The G37S substitution occurs at a residue which is absolutely conserved from bacteria to humans. Hs, *Homo sapiens*; Cf, *Canis familiaris*; Mm, *Mus musculus*; Rn, *Rattus norvegicus*; Ce, *Caenorhabditis elegans*; Sp, *Schizosaccharomyces pombe*; Sc, *Saccharomyces cerevisiae*; Ec, *Escherichia coli*; Ph, *Pyrococcus horikoshii*. **(d)** Electropherogram of *RNASEH2A* mutation. **(e)** Predicted tertiary structure of *RNASEH2A* catalytic site modeled on solved crystal structures of type 2 RNase H proteins. The substituted Gly37 residue (red) lies in close proximity to the active site (amino acid side chains of active site residues in green) and putative substrate-binding residues (light blue)³⁰.

major phylogenetic divisions (archaea, eubacteria, eukaryotes; Fig. 3c). Structurally, it lies at a turn just at the end of the first β -sheet, in the floor of the predicted substrate-binding cleft, close to the RNase H catalytic site (Fig. 3e)³⁰. We did not detect the mutation in 178 European control alleles, and both parents were heterozygous for the mutation.

RNASEH2A, B and C form a complex possessing RNase H2 activity

Homology with the *S. cerevisiae* Rnh2Ap-Rnh2Bp-Rnh2Cp complex predicts that their human counterparts AGS4, AGS2 and AGS3 might form a protein complex possessing RNase H2 activity (Fig. 4a). To address this prediction, we cloned the three human genes using the Gateway system into epitope-tagged mammalian expression vectors (pCGT-Dest (T7) and pCDNA3.1mycis-Dest) and transiently coexpressed all three constructs in HEK293T cells. Immunoprecipitation with antibodies to myc against C-terminally tagged *RNASEH2A*-myc pulled down N-terminally T7-tagged *RNASEH2C* and T7-tagged *RNASEH2B* (Fig. 4b), confirming that the three subunits interact *in vitro*.

We adapted a previously described fluorometric assay³¹ to test this complex for enzymatic activity. We annealed fluorescein-labeled oligonucleotides to a complementary DABCYL-labeled DNA oligonucleotide that quenched fluorescence (Fig. 4c). Enzymatic cleavage of the fluorescein oligonucleotide resulted in release of the fluorescein from the adjacent quencher molecule, generating a fluorescent signal (Fig. 4c).

Oligonucleotide (oligo) C, a substrate degradable by any ribonuclease H, is a 3'-fluorescein-tagged RNA oligonucleotide hybridized with a complementary 5'-labeled DABCYL DNA oligonucleotide. Oligo B, a substrate degradable only by type II ribonuclease H, is a DNA duplex with a single ribonucleotide at position 15 in the oligonucleotide strand that is 3'-labeled with fluorescein. Oligo A is a RNA:DNA hybrid like oligo C but is enzymatically resistant, as the 3'-fluorescein-labeled oligonucleotide has been synthesized with 2'-O-methyl RNA nucleotides.

We found that 'oligo C', an RNA:DNA duplex, was efficiently cleaved by the immunoprecipitated complex, confirming that this complex exhibits RNase H activity (Fig. 4d). Moreover, the immunoprecipitated complex possessed the requisite 'type 2' RNase H activity, as it recognized a single ribonucleotide embedded in a DNA-DNA duplex, and it efficiently cleaved 'oligo B', in contrast to *E. coli* RNase H, a 'type 1' RNase H. As expected, 'oligo A' was not

Notably, *RNASEH2C*, encoding a protein that biochemically copurifies with human *RNASEH2A* (ref. 23), lies within this critical interval (Fig. 2). By sequencing *RNASEH2C*, we then identified nonconservative missense mutations in all six *AGS3* families (Table 2 and Fig. 2). A homozygous mutation at codon 69 (resulting in the amino acid substitution R69W) was present in all affected individuals from the five Pakistani families that shared the apparent ancestral haplotype. A different homozygous mutation, resulting in K143I, was present in the Bangladeshi family. These mutations segregated with the disease within the families, and we did not detect either mutation in at least 172 Asian control alleles. RT-PCR expression profiling of *RNASEH2C* (data not shown) demonstrated a similar widespread expression pattern to *RNASEH2B*.

Concurrently, we established that *RNASEH2C* is the human ortholog of *S. cerevisiae* Rnh2Cp, a second subunit of yeast RNase H2 (ref. 21). A database search with a *Kluyveromyces waltii* open reading frame identifies both human *RNASEH2C* and *S. cerevisiae* Rnh2Cp within four search iterations (Supplementary Fig. 3). As we had now implicated two members of the RNase H2 complex in AGS, we then investigated the enzymatic subunit, *RNASEH2A*.

***RNASEH2A* is the *AGS4* gene**

Although *RNASEH2A* on 19p13.13 does not colocalize with any genetically mapped *AGS* loci, a review of our SNP array genome scan data identified a small region of homozygosity likely encompassing *RNASEH2A* in a previously described consanguineous *AGS* family of Spanish ancestry²⁹ (Fig. 3a). Sequencing in the two affected children from this family identified a single homozygous 109G→A transition, resulting in a G37S nonconservative missense substitution. (Fig. 3b,d). This amino acid is absolutely conserved across all three

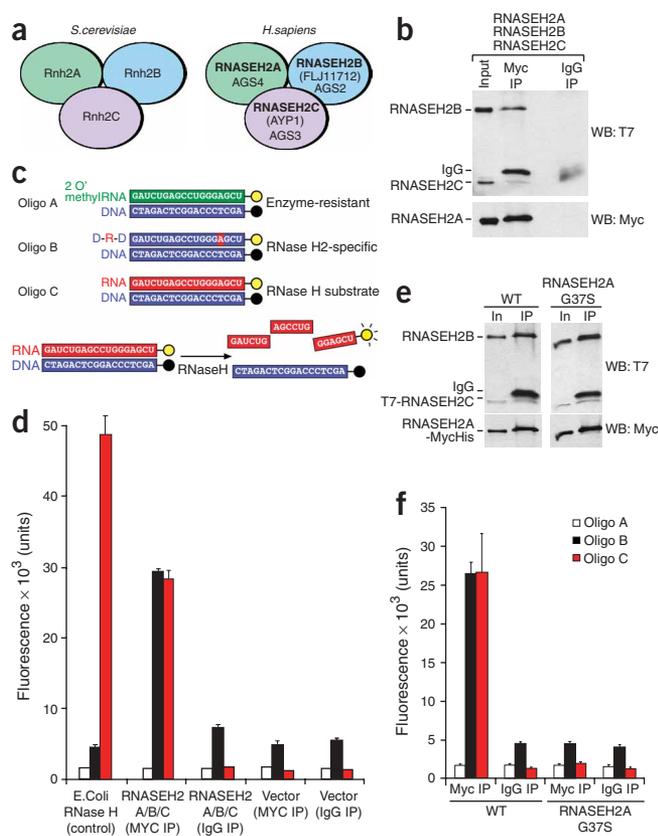


Figure 4 Human RNASEH2B, RNASEH2C and RNASEH2A form an enzymatically active type II ribonuclease H complex when expressed in mammalian cells, and mutation in *RNASEH2A* reduces ribonuclease H activity. **(a)** Schematic representation of the proposed human RNASEH2 complex and its *S. cerevisiae* counterpart. **(b)** T7 epitope-tagged RNASEH2B and RNASEH2C can be coimmunoprecipitated (IP) with myc-tagged RNASEH2A. **(c,d)** The RNASEH2B/RNASEH2C/RNASEH2A complex shows ribonuclease H activity. **(c)** Three oligonucleotide heteroduplexes were used to assay for enzymatic activity: oligo C, a substrate degradable by any ribonuclease H; oligo B, a substrate degradable only by type II ribonuclease H; and oligo A, an RNA:DNA hybrid like oligo C that is enzymatically resistant. **(d)** Type II ribonuclease H activity can be immunoprecipitated from HEK293T extracts containing epitope-tagged RNASEH2A ('A'), RNASEH2B ('B') and RNASEH2C ('C'). Myc IP was performed with mouse antibody to myc; 'IgG IP' indicates control immunoprecipitation performed with normal mouse IgG immunoglobulin. 'Vector' indicates cells transfected with empty pCGT-Dest and pcDNA3.1mychis vectors. Error bars: s.e.m. **(e)** Immunoprecipitation of the RNase H2 complex from HEK293T cells cotransfected with a mutant form of the RNASEH2A protein, in combination with wild-type tagged RNASEH2B and RNASEH2C, showing that the RNASEH2A G37S substitution does not seem to disrupt complex formation. In, input; IP, immunoprecipitate; WT, wild-type. **(f)** Mutation in the RNASEH2A reduces enzyme activity, as shown by fluorometric RNase H assay of the same immunoprecipitated complexes in **e**. Error bars: s.e.m.

cleaved, as it was synthesized using nuclease-resistant 2'-O-methyl RNA chemistry.

Mutation in *RNASEH2A* reduces enzymatic activity

We introduced the RNASEH2A G37S substitution using site-directed mutagenesis and coexpressed it with the other complex members by transient transfection of HEK293T cells. We performed immunoprecipitation to assess its effect on complex stability and to assay the complex's enzymatic activity (Fig. 4e,f). These experiments demonstrated that the RNASEH2A G37S substitution had no effect on complex stability. Rather, as expected from its proximity to catalytic and substrate-binding sites, it markedly reduced enzymatic activity (Fig. 4f).

DISCUSSION

Here we report the identification of three genes for Aicardi-Goutières syndrome. We establish that all three gene products are components of the ribonuclease H type 2 enzymatic complex. This represents the first demonstration that the mammalian RNase H2 enzyme contains three subunits that are orthologous to their yeast counterparts. Subsequent functional studies support our molecular data and homology predictions by showing that a disease-causing amino acid change in the complex can reduce enzymatic activity *in vitro*. The *AGS2*, *AGS3* and *AGS4* genes are therefore now designated *RNASEH2B* (formerly known as *FLJ11712*), *RNASEH2C* (formerly known as *AYP1*) and *RNASEH2A*, respectively.

Other pediatric disorders showing phenotypic overlap with AGS have been described^{32–35}. With the identification of the *AGS2*, *AGS3* and *AGS4* genes, it is now possible to address the phenotypic spectrum of AGS and clarify the relationship of AGS to these other diseases. Notably, the RNASEH2A G37S substitution seen in family F39

resulted in a phenotype that could be described equally well as either pseudo-TORCH or AGS²⁹. The finding of a heterozygote in the controls for the RNASEH2B A177T substitution also suggests that mutations in the RNase H2 genes may be a more frequent occurrence than previously recognized. However, further work is needed to accurately establish the carrier frequency of this condition in the population at large and also to functionally confirm the pathogenic nature of this particular mutation.

Cellular consequences of mutations in RNase H2 complex genes

RNase H2 is the major source of ribonuclease H activity in the cell^{23,25}, and reduced activity of this enzyme is likely to affect cellular processes dependent on RNA-DNA hybrid metabolism. Precisely how RNaseH2 deficiency leads to a cellular and immunological response that mimics congenital infection remains to be determined.

RNase H2 seems to have a role in the removal of Okazaki fragment RNA primers during lagging strand DNA replication^{24,26}. Although deletion of *S. cerevisiae Rnh2A* does not affect viability, it does result in increased sensitivity to hydroxyurea, indicating that RNase H2 mutations in AGS could be deleterious during episodes of replication stress²⁴. Also, unlike type 1 RNase H, RNase H2 can recognize single ribonucleotides embedded in DNA:DNA duplexes^{21,23}. This activity could be important in recognition and removal of inappropriately incorporated ribonucleotides from genomic DNA²⁵. Such misincorporation might occur more frequently in circumstances of depleted dNTPs and might provide an alternative explanation for the hydroxyurea sensitivity of *Rnh2A* mutants²⁴. Although these functions suggest a role for RNase H2 in genomic integrity, it is noteworthy that individuals with AGS do not show increased cancer susceptibility.

Ribonuclease H2 and the 'viral infection' phenotype

Reverse transcription is an essential process for retrovirus replication³⁶, and DNA-RNA hybrids formed during this process could be susceptible to degradation by endogenous RNase H. Therefore, abrogation of RNase H2 function might impair host antiviral defenses through reduced destruction of retroviral DNA-RNA hybrids. However, despite extensive study, viral pathogens have not been identified in individuals with AGS⁸, and it seems more likely that endogenous

effects of AGS mutations cause the viral infection-like phenotype. dsRNA and dsDNA are known activators of innate immunity and stimulate type I interferon production³⁷. Therefore, a possible disease mechanism arising from reduced RNase H2 function could involve raised levels of endogenous RNA-DNA hybrids stimulating interferon α production by mechanisms similar to those for dsRNA and dsDNA. Consequent overproduction of interferon α in the central nervous system may then explain many of the neuropathological features of AGS, as shown by a transgenic mouse model in which interferon α has been chronically expressed in glial cells³⁸.

Summary

We show that mutations in any of the genes encoding the three subunits of the RNase H2 enzyme complex cause AGS and may do so as a result of reduced enzyme function. This implicates a cellular process involving RNA-DNA hybrids as the pathogenic basis for the disorder. The association of ribonuclease H with a 'viral infection' phenotype indicates that further characterization of this nuclease should provide insights into host antiviral response.

METHODS

Individuals affected with AGS. All affected individuals included in the study fulfilled diagnostic criteria for AGS, with neurological features of an early-onset encephalopathy, negative investigations for common prenatal infections, intracranial calcification in a typical distribution, a CSF lymphocytosis >5 cells/mm³ and/or >2 international units/ml of interferon- α in the CSF (also see **Supplementary Note** and **Supplementary Fig. 5** online). With informed consent, blood samples were obtained from affected children, their parents, and unaffected siblings. Genomic DNA was extracted from peripheral blood leukocytes by standard methods. The study was approved by the Leeds Health Authority/United Teaching Hospitals National Health Service Trust Research Ethics Committee and by the Scottish Multicentre Research Ethics Committee (04:MRE00/19).

Genotyping and linkage analysis. Genome-wide scans by SNP array were performed using Affymetrix Human Mapping10K Xba142 2.0 GeneChips by the UK MRC Geneservice. High-density genotyping of AGS2 and AGS3 loci was performed as previously described³⁹ using established microsatellite markers from the Marshfield linkage maps and new microsatellites identified from the UCSC Human Genome Browser sequence (May 2004 freeze). Linkage analysis was performed using GENEHUNTER (version 2.0 β)⁴⁰ under a model of autosomal recessive inheritance with full penetrance, with a disease allele frequency estimated at 1 in 1,000 and assuming equal marker allele frequencies.

Bioinformatics. Database searches were conducted with PSI-BLAST²⁰ using the non-redundant protein sequence database and an *E*-value inclusion threshold of 2×10^{-3} . Alignments were viewed using CHROMA⁴¹.

Comparative protein structure modeling was performed using SWISS-MODEL⁴² with default settings and accuracy of the RNASEH2A protein (NP_006388) predicted structure validated by WHAT_CHECK⁴³. Four homologous archaeal RNase HII proteins (IuaxA, Iio2A, Ix1pA and IekeB) of known structure (demonstrating 35–40% sequence identity with NP_006388) were used as best available templates for the prediction. The active site and predicted substrate contact residues³⁰ were annotated using VMD (v1.8.3)⁴⁴.

Mutation detection. Primers were designed to amplify the coding exons of RNASEH2B, RNASEH2C and RNASEH2A (primer sequences listed in **Supplementary Table 1** online). Purified PCR amplification products were sequenced using dye-terminator chemistry and electrophoresed on ABI 3700 (Applied Biosystems) or MegaBace500 (Amersham Pharmacia) capillary sequencers. Mutation analysis was performed using Mutation Surveyor (Softgenetics). Anonymized control samples were screened by sequencing. Controls were matched with the ethnic origin of the mutations for the respective genes. Mutations in RNASEH2B were from diverse ethnic groups, although Europeans predominated, and therefore European controls were screened.

Vector construction. The Gateway Vector system (Invitrogen) was used to construct mammalian expression vectors. The coding sequences of RNASEH2B and RNASEH2A were amplified from plasmid clones (clones CS0DF031YM15 (GenBank accession number CR602872) from Invitrogen and clones IRAUp969G0361D (GenBank accession number BC011748) from RZPD) and were cloned into pDONR221 (Invitrogen). RNASEH2C was purchased as a ready-made pENTRY vector (clone IOH27907, Human Ultimate Full ORF Gateway Shuttle Clone, NM_032193, Invitrogen). pcDNA3.1mychis-Dest and pCGT-Dest were constructed using the Gateway Vector conversion system by insertion of Gateway reading frame cassettes into the multicloning sites of pcDNA3.1mychis (Invitrogen) and pCGT⁴⁵. Site-directed mutagenesis was performed on the RNASEH2A pENTRY clone using the Stratagene Quikchange kit according to the manufacturer's instructions.

Immunoprecipitation and western blotting. HEK293T cells were transiently cotransfected with 1 μ g of each construct using Lipofectamine (Invitrogen) according to manufacturer's instructions. After 24 hours, cell were lysed in 50 mM Tris (pH 7.8), 280 mM NaCl, 0.5% NP40, 0.2 mM EDTA, 0.2 mM EGTA, 10% glycerol (vol/vol), 0.1 mM sodium orthovanadate, 1 μ M DTT and 1 μ M PMSF for 10 min at 4 °C. Lysed cells were then diluted 1:1 with 20 mM HEPES (pH 7.9), 10 mM KCl, 1 mM EGTA, 10% glycerol and 0.1 mM sodium orthovanadate buffer for an additional 10 min, and extracts were cleared by centrifugation (15,800g, 10 min, 4 °C). We immunoprecipitated 500 μ g of protein lysate using Protein A/G PLUS agarose (Santa Cruz), following the manufacturer's protocol, using 1 μ g of mouse antibody to myc (clone 9B11, Cell Signalling) or 1 μ g of mouse IgG (Santa Cruz). Protein blots were performed using mouse monoclonal antibody to myc at 1:1,000 and mouse monoclonal antibody to T7 (Novagen) at 1:5,000.

RNase H assays. We annealed 10 μ M oligonucleotides (Eurogentec) in 60 mM KCl, 50 mM Tris-HCl (pH 8) by denaturation at 95 °C for 5 min and then gradual cooling to room temperature (18 °C). Fluorometric RNase H assays were performed in a 100- μ l volume of 60 mM KCl, 50 mM Tris-HCl (pH 8), 10 mM MgCl₂ and 0.25 μ M oligonucleotide duplex in 96-well flat-bottomed plates at 37 °C for 3 h in an orbital shaker at 60 rpm. One-tenth the volume of each immunoprecipitate was used in each reaction, and as a positive control, 2.5 units of *E. coli* RNase H (Invitrogen) was used. 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.

Accession codes. GenBank: RNASEH2B, NM_024570; RNASEH2C, AF312034; RNASEH2A, NM_006397.

URLs. Marshfield linkage map: <http://research.marshfieldclinic.org/genetics>; UCSC Human Genome Browser: <http://genome.ucsc.edu/>; PSI-BLAST: <http://www.ncbi.nlm.nih.gov/blast/>.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank the families and their clinicians for their participation in this study; G. Taylor, S. Farrington and C. Hayward for contributing control samples; A. Diamond for advice on Mutation Surveyor; S. McKay and the MRC HGU core sequencing service for advice and technical support; D. Stuart for preparation of illustrations; P. Hohenstein and N. Gilbert for assistance with reagents; V. Van Heyningen, J. Sanford, D. Fitzpatrick, W. Bickmore, B. Vernay, A. Wright and N. Hastie for discussions and comments; F.B. Longo and the International Aicardi-Goutières syndrome Association for their encouragement and K. Norton, L. Cervero and G. Pitelet for their help clinically. This work was supported by the MRC, the Fondazione Cariplo, The Leeds Teaching Hospitals Charitable Foundation and the West Riding Medical Research Trust. A.P.J. is an MRC Clinician Scientist, and A.P.J. and CPP are funded by the MRC.

AUTHOR CONTRIBUTIONS

Y.J.C. and R.P. performed the microsatellite genotyping; A.P.J. performed linkage analysis and A.G., A.L., B.H., R.P. and A.P.J. performed mutation screening and sequencing of controls. A.L. and A.P.J. made the vector constructs and designed the enzyme assays. A.L. and E.G. performed the immunoprecipitations and enzyme assays. C.P.P. established orthology between RNASEH2B and Rnh2Bp and between RNASEH2C and Rnh2Cp, and C.S. performed the structural modeling.

A.P.J. wrote the paper with Y.J.C., D.T.B. and C.P.P. Y.J.C. curated the clinical samples and data. All other authors provided clinical samples and data.

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests (see the *Nature Genetics* website for details).

Published online at <http://www.nature.com/naturegenetics>

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>

- Aicardi, J. & Goutières, F. A progressive familial encephalopathy in infancy with calcifications of the basal ganglia and chronic cerebrospinal fluid lymphocytosis. *Ann. Neurol.* **15**, 49–54 (1984).
- Goutières, F. Aicardi-Goutières syndrome. *Brain Dev.* **27**, 201–206 (2005).
- Goutières, F., Aicardi, J., Barth, P.G. & Lebon, P. Aicardi-Goutières syndrome: an update and results of interferon-alpha studies. *Ann. Neurol.* **44**, 900–907 (1998).
- Bale, J.F., Jr., Bray, P.F. & Bell, W.E. Neuroradiographic abnormalities in congenital cytomegalovirus infection. *Pediatr. Neurol.* **1**, 42–47 (1985).
- Numazaki, K. & Fujikawa, T. Intracranial calcification with congenital rubella syndrome in a mother with serologic immunity. *J. Child Neurol.* **18**, 296–297 (2003).
- Mitchell, W. Neurological and developmental effects of HIV and AIDS in children and adolescents. *Ment. Retard. Dev. Disabil. Res. Rev.* **7**, 211–216 (2001).
- Belman, A.L. *et al.* AIDS: calcification of the basal ganglia in infants and children. *Neurology* **36**, 1192–1199 (1986).
- Lebon, P., Meritet, J.F., Krivine, A. & Rozenberg, F. Interferon and Aicardi-Goutières syndrome. *Eur. J. Paediatr. Neurol.* **6**(Suppl.), A47–A53 (2002).
- Krivine, A. *et al.* Measuring HIV-1 RNA and interferon-alpha in the cerebrospinal fluid of AIDS patients: insights into the pathogenesis of AIDS Dementia Complex. *J. Neurovirology* **5**, 500–506 (1999).
- Lebon, P., Ponsot, G., Aicardi, J., Goutières, F. & Arthus, M. Early intrathecal synthesis of interferon in herpes encephalitis. *Biomedicine* **31**, 267–271 (1979).
- Dussaix, E., Lebon, P., Ponsot, G., Huault, G. & Tardieu, M. Intrathecal synthesis of different alpha-interferons in patients with various neurological diseases. *Acta Neurol. Scand.* **71**, 504–509 (1985).
- Kumar, D., Rittey, C., Cameron, A.H. & Variend, S. Recognizable inherited syndrome of progressive central nervous system degeneration and generalized intracranial calcification with overlapping phenotype of the syndrome of Aicardi and Goutières. *Am. J. Med. Genet.* **75**, 508–515 (1998).
- Crow, Y.J. *et al.* Aicardi-Goutières syndrome displays genetic heterogeneity with one locus (AGS1) on chromosome 3p21. *Am. J. Hum. Genet.* **67**, 213–221 (2000).
- Crow, Y.J. *et al.* Cree encephalitis is allelic with Aicardi-Goutières syndrome: implications for the pathogenesis of disorders of interferon alpha metabolism. *J. Med. Genet.* **40**, 183–187 (2003).
- Ali, M. *et al.* A second locus for Aicardi-Goutières syndrome at chromosome 13q14–21. *J. Med. Genet.* **43**, 444–450 (2006).
- Broman, K.W. & Weber, J.L. Long homozygous chromosomal segments in reference families from the Centre d'Etude du Polymorphisme Humain. *Am. J. Hum. Genet.* **65**, 1493–1500 (1999).
- Gibson, J., Morton, N.E. & Collins, A. Extended tracts of homozygosity in outbred human populations. *Hum. Mol. Genet.* **15**, 789–795 (2006).
- Lander, E.S. & Botstein, D. Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. *Science* **236**, 1567–1570 (1987).
- Henikoff, S. & Henikoff, J.G. Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. USA* **89**, 10915–10919 (1992).
- Altschul, S.F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
- Jeong, H.S., Backlund, P.S., Chen, H.C., Karavanov, A.A. & Crouch, R.J. RNase H2 of *Saccharomyces cerevisiae* is a complex of three proteins. *Nucleic Acids Res.* **32**, 407–414 (2004).
- Ohtani, N. *et al.* Identification of the genes encoding Mn²⁺-dependent RNase HIII and Mg²⁺-dependent RNase HIII from *Bacillus subtilis*: classification of RNases H into three families. *Biochemistry* **38**, 605–618 (1999).
- Frank, P., Braunschöfer-Reiter, C., Wintersberger, U., Grimm, R. & Busen, W. Cloning of the cDNA encoding the large subunit of human RNase H1, a homologue of the prokaryotic RNase HIII. *Proc. Natl. Acad. Sci. USA* **95**, 12872–12877 (1998).
- Arudchandran, A. *et al.* The absence of ribonuclease H1 or H2 alters the sensitivity of *Saccharomyces cerevisiae* to hydroxyurea, caffeine and ethyl methanesulphonate: implications for roles of RNases H in DNA replication and repair. *Genes Cells* **5**, 789–802 (2000).
- Eder, P.S. & Walder, J.A. Ribonuclease H from K562 human erythroleukemia cells. Purification, characterization, and substrate specificity. *J. Biol. Chem.* **266**, 6472–6479 (1991).
- Qiu, J., Qian, Y., Frank, P., Wintersberger, U. & Shen, B. *Saccharomyces cerevisiae* RNase H(35) functions in RNA primer removal during lagging-strand DNA synthesis, most efficiently in cooperation with Rad27 nuclease. *Mol. Cell. Biol.* **19**, 8361–8371 (1999).
- Eder, P.S., Walder, R.Y. & Walder, J.A. Substrate specificity of human RNase H1 and its role in excision repair of ribose residues misincorporated in DNA. *Biochimie* **75**, 123–126 (1993).
- Rydberg, B. & Game, J. Excision of misincorporated ribonucleotides in DNA by RNase H (type 2) and FEN-1 in cell-free extracts. *Proc. Natl. Acad. Sci. USA* **99**, 16654–16659 (2002).
- Sanchis, A. *et al.* Genetic syndromes mimic congenital infections. *J. Pediatr.* **146**, 701–705 (2005).
- Chapados, B.R. *et al.* Structural biochemistry of a type 2 RNase H: RNA primer recognition and removal during DNA replication. *J. Mol. Biol.* **307**, 541–556 (2001).
- Parniak, M.A., Min, K.L., Budihas, S.R., Le Grice, S.F. & Beutler, J.A. A fluorescence-based high-throughput screening assay for inhibitors of human immunodeficiency virus-1 reverse transcriptase-associated ribonuclease H activity. *Anal. Biochem.* **322**, 33–39 (2003).
- Baraitser, M., Brett, E.M. & Piesowicz, A.T. Microcephaly and intracranial calcification in two brothers. *J. Med. Genet.* **20**, 210–212 (1983).
- Reardon, W. *et al.* Autosomal recessive congenital intrauterine infection-like syndrome of microcephaly, intracranial calcification, and CNS disease. *Am. J. Med. Genet.* **52**, 58–65 (1994).
- Schwarz, K.B., Ferrie, C.D. & Woods, C.G. Two siblings with a new Aicardi-Goutières-like syndrome. *Dev. Med. Child Neurol.* **44**, 422–425 (2002).
- Gardner, R.J. *et al.* Severe fetal brain dysgenesis with focal calcification. *Prenat. Diagn.* **25**, 362–364 (2005).
- Coffin, J.M., Hughes, S.H. & Varmus, H.E. *Retroviruses* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1997).
- Kawai, T. & Akira, S. Innate immune recognition of viral infection. *Nat. Immunol.* **7**, 131–137 (2006).
- Campbell, I.L. *et al.* Structural and functional neuropathology in transgenic mice with CNS expression of IFN-alpha. *Brain Res.* **835**, 46–61 (1999).
- Jackson, A.P. *et al.* Primary autosomal recessive microcephaly (MCPH1) maps to chromosome 8p22-pter. *Am. J. Hum. Genet.* **63**, 541–546 (1998).
- Kruglyak, L., Daly, M.J., Reeve-Daly, M.P. & Lander, E.S. Parametric and non-parametric linkage analysis: A unified multipoint approach. *Am. J. Hum. Genet.* **58**, 1347–1363 (1996).
- Goodstadt, L. & Ponting, C.P. CHROMA: consensus-based colouring of multiple alignments for publication. *Bioinformatics* **17**, 845–846 (2001).
- Schwede, T., Kopp, J., Guex, N. & Peitsch, M.C. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res.* **31**, 3381–3385 (2003).
- Hoof, R.W., Vriend, G., Sander, C. & Abola, E.E. Errors in protein structures. *Nature* **381**, 272 (1996).
- Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38–27–28 (1996).
- Van Aelst, L., Joneson, T. & Bar-Sagi, D. Identification of a novel Rac1-interacting protein involved in membrane ruffling. *EMBO J.* **15**, 3778–3786 (1996).