The evolving genetic and pathophysiological spectrum of migraine

Vries, B. de

5.1

C-terminal truncations in human 3′-5′ DNA exonuclease TREX1 cause autosomal dominant retinal vasculopathy with cerebral leukodystrophy

Anna Richards[1,*], Arn M J M van den Maagdenberg[2,3,*], Joanna C Jen[4,*], David Kavanagh[1,*], Paula Bertram[1], Dirk Spitzer[1], M Kathryn Liszewski[1], Maria-Louise Barilla-LaBarca[5], Gisela M Terwindt[3], Yumi Kasai[6], Mike McLellan[6], Mark Gilbert Grand[7], Kaate R J Vanmolkot[2], Boukje de Vries[2], Jijun Wan[4], Michael J Kane[4], Hafsa Mamsa[4], Ruth Schäfer[4], Anine H Stam[3], Joost Haan[3], Paulus T V M de Jong[8–10], Caroline W Storimans[11], Mary J van Schooneveld[12], Jendo A Oosterhuis[13], Andreas Gschwendter[14], Martin Dichgans[14], Katya E Kotschet[15], Suzanne Hodgkinson[16], Todd A Hardy[17], Martin B Delatycki[18,19], Rula A Hajj-Ali[20], Parul H Kothari[1], Stanley F Nelson[21], Rune R Frants[2], Robert W Baloh[4], Michel D Ferrari[3] & John P Atkinson[1]

[1]Dept. of Medicine, Division of Rheumatology, Washington University School of Medicine, St. Louis, Missouri 63110, USA. [2]Dept. of Human Genetics, and [3]Dept. of Neurology, Leiden University Medical Centre, 2300 RC Leiden, The Netherlands. [4]Dept. of Neurology, University of California at Los Angeles, Los Angeles, California 90095, USA. [5]Dept. of Medicine, Division of Rheumatology, North Shore Long-Island Jewish Health System, Lake Success, New York 11030, USA. [6]Genome Sequencing Center, and [7]Department of Ophthalmology, Washington University School of Medicine, St. Louis, Missouri 63110, USA. [8]Dept. of Ophthalmogenetics, Netherlands Institute for Neuroscience, Royal Netherlands Academy of Arts and Sciences, 1000 GC Amsterdam, The Netherlands. [9]Dept. of Ophthalmology, Academic Medical Centre, 1100 DD Amsterdam, The Netherlands. [10]Dept. of Epidemiology and Biostatistics, Erasmus Medical Centre, 3000 CA Rotterdam, The Netherlands. [11]Meander Medical Centre, 3800 BM Amersfoort, The Netherlands. [12]Dept. of Ophthalmology, University Medical Centre, 3508 GA Utrecht, The Netherlands. [13]Dept. of Ophthalmology, Leiden University Medical Centre, 2300 RC Leiden, The Netherlands. [14]Dept. of Neurology, Klinikum Grosshadern, Universität München, D-81377 München, Germany. [15]Dept. of Neurology, Monash Medical Centre, Clayton, Victoria 3168, Australia. [16]Dept. of Neurology, Liverpool Hospital, Liverpool, New South Wales 2170, Australia. [17]Dept. of Neurology, Concord Repatriation General Hospital, Concord, New South Wales 2139, Australia. [18]Bruce Lefroy Centre for Genetic Health Research, Murdoch Childrens Research Institute, and [19]Dept. of Paediatrics, University of Melbourne, Royal Children's Hospital, Parkville, Victoria 3052, Australia. [20]Dept. of Rheumatic and Immunologic Diseases, Cleveland Clinic Foundation, Cleveland, Ohio 44195, USA. [21]Dept. of Human Genetics, University of California at Los Angeles, Los Angeles, California 90095, USA. [*]These authors contributed equally to this paper

Nature Genetics 2007;39:1068-1070

Autosomal dominant retinal vasculopathy with cerebral leukodystrophy is a microvascular endotheliopathy with middle-age onset. In nine families, we identified heterozygous C-terminal frameshift mutations in *TREX1*, which encodes a 3′-5′ exonuclease. These truncated proteins retain exonuclease activity but lose normal perinuclear localization. These data have implications for the maintenance of vascular integrity in the degenerative cerebral microangiopathies leading to stroke and dementias.

We have previously described three families sharing common features of retinal and cerebral dysfunction. Visual loss, stroke and dementia begin in middle age, and death occurs in most families 5 to 10 years later. These diseases map to 3p21.1–p21.3 (ref. 1) and are called cerebroretinal vasculopathy (CRV)[2], hereditary vascular retinopathy (HVR)[3,4] and hereditary endotheliopathy, retinopathy, nephropathy and stroke (HERNS)[5]. We now designate these illnesses as autosomal dominant retinal vasculopathy with cerebral leukodystrophy (RVCL) (OMIM 192315). The neurovascular syndrome features a progressive loss of visual acuity secondary to retinal vasculopathy, in combination with a more variable neurological picture[1–7]. In a subset of affected individuals, systemic vascular involvement is evidenced by Raynaud's phenomenon and mild liver (micronodular cirrhosis)[2,5] and kidney (glomerular) dysfunction[5].

This retinal vasculopathy is characterized by telangiectasias, microaneurysms and retinal capillary obliteration starting in the macula. Diseased cerebral white matter has prominent small infarcts that often coalesce to pseudotumors. Neuroimaging studies demonstrate contrast-enhancing lesions in the white matter of the cerebrum and cerebellum. Histopathology shows ischemic necrosis with minimal inflammation and small blood vessels occluded with fibrin[5]. The white matter lesions resemble post-radiation vascular damage[2]. Ultrastructural studies of capillaries show a distinctive, multilamellar subendothelial basement membrane[5].

By combining haplotypes in the three RVCL families, we narrowed the disease gene to a 3-cM region between markers D3S1578 and D3S3564 that encompassed ~10 Mb, containing over 120 candidate genes[1]. We then sequenced the full coding region and intron-exon boundaries of 33 candidate genes within this region (Supplementary Table 1).

Here we report the identification of mutations in *TREX1* (NM_033627), encoding DNA-specific 3′ to 5′ exonuclease DNase III. In the CRV[2] and HVR3,[4] pedigrees, a heterozygous 1-bp insertion (3688_3689insG) leads to V235fs and a consequent premature stop. In HERNS[5], a heterozygous 4-bp insertion (3727_3730dupGTCA) results in a frameshift at T249 (Fig. 1a,b).

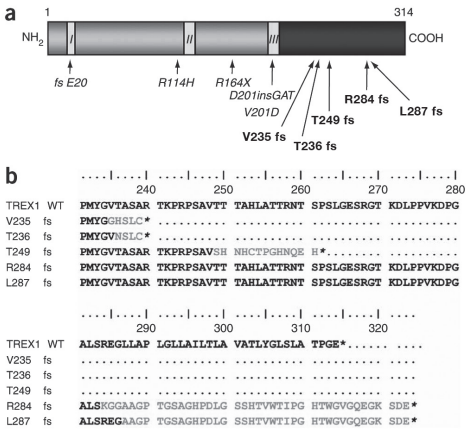


Figure 1 *Diagram of TREX1 protein.* **a)** *TREX1 has three exonuclease domains. Mutations in italics are associated with AGS[13], and those in boldface at the C terminus are associated with RVCL.* **b)** *Comparison of the amino acid sequence of the C terminus of wild-type (WT) TREX1 with RVCL associated mutations. The abnormal sequence introduced by the frameshifts is depicted in gray.*

Next, we examined six families with putative RVCL (Supplementary Table 2)[2,6,7]. In each, we identified frameshift mutations affecting the C terminus of *TREX1*. In three, the alteration was V235fs, the same as that in the CRV and HVR pedigrees. Haplotype analysis suggests that they are not related (data not shown). We did not detect any of the mutations in panels of chromosomes matched by ancestry or location (Supplementary Methods). In the CRV and HERNS families, all affected individuals over the age of 60 (but none of the unaffected individuals over the age of 60) carried a *TREX1* mutation (100% penetrance). In the HVR[3,4] family, 10 of the 11 mutation carriers over 60 years of age have retinopathy.

TREX1 (DNase III) is a DNA-specific 3′ to 5′ exonuclease ubiquitously expressed in mammalian cells[8–10]. It is thought to function as a homodimer, with a preference for single-stranded DNA and mispaired 3′ termini[8]. *TREX1* is a part of the SET complex[11] that normally resides in the cytoplasm but translocates to the nucleus in response to oxidative DNA damage[12].

Recently, homozygous mutations in *TREX1* have been reported to cause Aicardi-Goutiere syndrome (AGS)[13]. AGS is a rare, familial, early-onset progressive encephalopathy featuring basal ganglia calcifications and cerebrospinal fluid lymphocytosis, mimicking congenital viral encephalitis[14]. Notably, mutations associated with AGS disrupt the enzymatic sites in *TREX1*. This loss of exonuclease function[13] (Fig. 1) is hypothesized to cause the accumulation of altered DNA that triggers a destructive autoimmune response[13]. No phenotype was reported for the heterozygous carriers of these mutations; however, a heterozygous mutation in *TREX1* causing familial chilblain lupus has been reported recently[15].

The distinctive clinical course and pathology of RVCL compared with AGS suggests separate disease mechanisms. The frameshift mutations observed in RVCL are downstream of the regions encoding the catalytic domains, whereas in AGS, homozygous mutations occur that alter exonuclease function. The heterozygous mutations observed in RVCL did not impair the enzymatic activity of TREX1 (Fig. 2a), in comparison with the R114H substitution in AGS[13].

To investigate how the RVCL TREX1 proteins differ from the wild-type, we performed expression studies using confocal microscopy on cells transfected with TREX1 tagged with a fluorescent protein (Fig. 2b and Supplementary Fig. 1). The wild-type TREX1 labeled with fluorescent protein (FP-TREX1) localized to the perinuclear region. In contrast, the TREX1 proteins FP-V235fs and FP-T249fs were diffusely distributed in the cytoplasm and the nucleus, as was the case for the fluorescent protein alone (Fig. 2b and Supplementary Videos 1–4 online). Protein blotting confirmed that the expressed proteins were of the correct size (Fig. 2c). These results suggest a perinuclear targeting signal within the C terminus of TREX1. Consequently, we generated a construct containing the C-terminal 106 amino acid residues of TREX1 (FP-C-106). This protein showed a perinuclear localization pattern identical to that of the wild-type TREX1 protein (Fig. 2b). The TREX1 protein containing amino acid change R114H, found in AGS, also had the same pattern as the wild-type protein. In contrast, the protein with the alteration closest to the C terminus of TREX1, FP-287fs, was diffusely distributed, like the other two truncated proteins (data not shown).

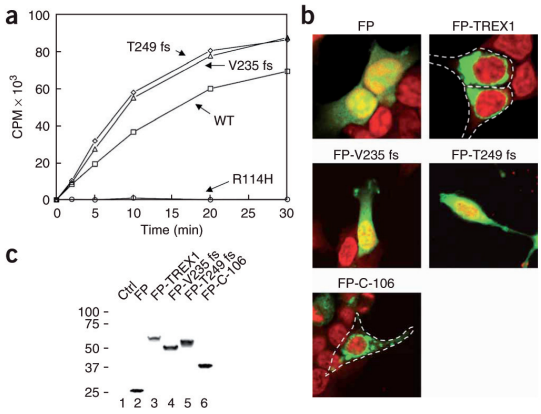


Figure 2 *Functional consequences of RVCL associated TREX1 mutations.* **a)** *Assessment of 3′-5′ exonuclease activity using equivalent amounts of purified recombinant proteins expressed in E. coli.* **b)** *Confocal microscopy of HEK293T cells showing transiently expressed fluorescent protein (FP)-tagged TREX1 proteins (green), TOPRO3 staining of nuclei (red) and overlay (yellow). Similar expression patterns were obtained for wild-type protein and for proteins derived from constructs containing mutations associated with AGS and RVCL in CHO, HL-60 and HeLa cells (data not shown).* **c)** *Protein blot analysis of untransfected cells (1) and cells transfected with enhanced yellow fluorescent protein (eYFP) (2), wild-type TREX1 (3), TREX1 mutants (4,5) and the C-terminal 106 amino acids (6), all linked to eYFP.*

The TREX1 proteins found in individuals with RVCL lack part of the C terminus. In haploinsufficient individuals, thismay prevent an interaction with the SET proteins and therefore may prevent formation of the SET complex. The SET complex is hypothesized to target DNA repair factors, including TREX1, to damaged DNA under conditions of oxidative stress[11,12]. Lack of sufficient TREX1 associated with the SET complex may result in failure of granzyme A–mediated cell death[12]. Alternatively, the dissemination of untethered TREX1 in the nucleus and cytoplasm may have detrimental effects, especially on endothelial cells.

The clinical syndromes in these families and the study of their mutations should deepen our understanding of exonuclease function, homeostasis of the endothelium and events leading to premature vascular aging. RVCL and cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) represent two examples of monogenic disease featuring a cerebral microangiopathy for which the genetic defects are now known and from which we can gain new insights into the origin of strokes and dementia.

We obtained consent from all participants in this study, and the study was approved by the Office for Protection of Research Subjects at UCLA and the Human Research Protection Office at Washington University School of Medicine.

Acknowledgements

We appreciate the cooperation of the participating families. We thank M. RBogacki, E. van den Boogerd, J. van Vark and S. Keradhmand-Kia. A.R. is a 2005/2006 Fulbright Distinguished Scholar. D.K. is a Kidney Research UK clinical training fellow. A.R. and D.K. are recipients of Peel Medical Trust Travel Fellowships. At Washington University in St. Louis, this study has been funded by the Center for Genome Sciences Pilot-Scale Sequencing Project Program and by the Danforth Foundation. The Netherlands Organization for Scientific Research (NWO) (Vici 918.56.602), the European Union “Eurohead” grant (LSHMCT- 2004-504837) and the Center of Medical System Biology established by the Netherlands Genomics Initiative/NWO supported the work in the Netherlands. US National Institutes of Health (NIH)/National Institute on Deafness and Other Communication Disorders (NIDCD) grant P50 DC02952 (R.W.B.), NIH/National Eye Institute grant R01 EY15311 and a Stein-Oppenheimer Award (J.C.J.) supported the work at University of California, Los Angeles. R.S. is the recipient of a scholarship from the German National Scholarship Foundation.

Competing interests statement

The authors declare no competing financial interests. Published online at http://www.nature.com/naturegenetics

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

References

1. Ophoff RA, DeYoung J, Service SK, et al (2001) Hereditary vascular retinopathy, cerebroretinal vasculopathy, and hereditary endotheliopathy with retinopathy, nephropathy, and stroke map to a single locus on chromosome 3p21.1-p21.3. *Am. J. Hum. Genet.* 69, 447–453.

2. Grand MG, Kaine J, Fulling K et al (1988) Cerebroretinal vasculopathy. A new hereditary syndrome. *Ophthalmology* 95, 649–659.

3. Storimans CW, Van Schooneveld MJ, Oosterhuis JA, Bos PJ (1991) A new autosomal dominant vascular retinopathy syndrome. *Eur. J. Ophthalmol* 1, 73–78.

4. Terwindt GM, Haan J, Ophoff RA et al (1998) Clinical and genetic analysis of a large Dutch family with autosomal dominant vascular retinopathy, migraine and Raynaud's phenomenon. *Brain* 121, 303–316.

5. Jen J, Cohen AH, Yue Q et al (1997) Hereditary endotheliopathy with retinopathy, nephropathy, and stroke (HERNS). *Neurology* 49, 1322–1330.

6. Cohn AC, Kotschet K, Veitch A et al (2005) Novel ophthalmological features in hereditary endotheliopathy with retinopathy, nephropathy and stroke syndrome. *Clin. Experiment. Ophthalmol.* 33, 181–183.

7. Weil, S, Reifenberger G, Dudel C et al (1999) Cerebroretinal vasculopathy mimicking a brain tumor: a case of a rare hereditary syndrome. *Neurology* 53, 629–631 (1999).

8. Mazur DJ & Perrino FW (2001) Excision of 3′ termini by the Trex1 and TREX2 3′→5′ exonucleases. Characterization of the recombinant proteins. *J. Biol. Chem.* 276, 17022–17029.

9. Mazur DJ & Perrino FW (1999) Identification and expression of the TREX1 and TREX2 cDNA sequences encoding mammalian 3′→5′ exonucleases. *J. Biol. Chem*. 274, 19655–19660.

10. Hoss M, Robins P, Naven TJ et al (1999) A human DNA editing enzyme homologous to the Escherichia coli DnaQ/MutD protein. *EMBO J*. 18, 3868–3875.

11. Chowdhury D, Beresford PJ, Zhu P et al (2006) The exonuclease TREX1 is in the SET complex and acts in concert with NM23-H1 to degrade DNA during granzyme A-mediated cell death. *Mol. Cell* 23, 133–142.

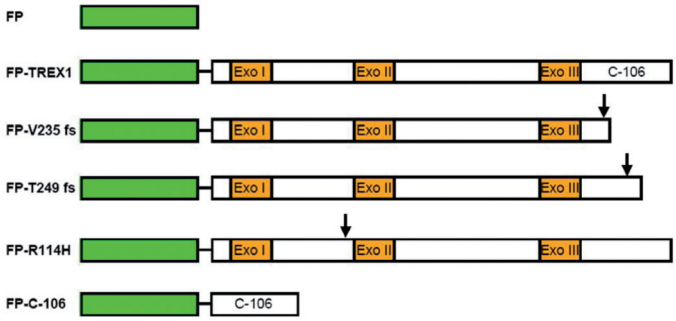
12. Martinvalet D, Zhu P & Lieberman J (2005) Granzyme A induces caspase-independent mitochondrial damage, a required first step for apoptosis. *Immunity* 22, 355–370.

13. Crow YJ, Hayward BE, Parmar R et al (2006) Mutations in the gene encoding the 3′-5′ DNA exonuclease TREX1 cause Aicardi-Goutières syndrome at the AGS1 locus. *Nat. Genet*. 38, 917–920.

14. Goutieres F (2005) Aicardi-Goutières syndrome. *Brain Dev*. 27, 201–206.

15. Rice G, Newman WG, Dean J et al (2007) Heterozygous mutations in TREX1 cause familial chilblain lupus and dominant Aicardi-Goutieres syndrome. *Am. J. Hum. Genet*. 80, 811–815.

Supplementary Material



Supplementary Fig. 1. *Schematic representation of FP constructs expressed in mammalian cells. The FP was cloned at the amino-terminus of TREX1 and mutants. A carboxyl-terminal segment of the last 106 amino acids of wild-type TREX1 was also prepared (C-106). The arrow indicate approximate sites of the mutations.*

Supplementary Table 1. *The 32 candidate genes sequenced prior to the discovery of TREX1 as the causative gene for RVCL.*

Gene	Name	Gene ID	OMIM
AMIGO3	Adhesion molecule with Ig-like domain 3	386724	N/A
ATRIP	ATR interacting protein	11277	606605
CACNA1D	Voltage-dependent L-type calcium channel subunit alpha-1D	776	114206
CACNA2D2	Calcium channel voltage dependent, alpha- 2/ Delta Subunit 2	9254	607082
CCR1	Chemokine (C-C motif) receptor 1	1230	601159
CCR2	Chemokine (C-C motif) receptor 2	1231	601267
CCR3	Chemokine (C-C motif) receptor 3	1232	601268
CCR9	Chemokine (C-C motif) receptor 9	10803	604738
CELSR3	Cadherin, EGF LAG seven-pass G-type receptor 3	1951	604264
CSPG5	Chondroitin sulfate proteoglycan 5	10675	606775
CTNNB1	Catenin (cadherin-associated protein), beta 1	1499	116806
CX3CR1	Chemokine (C-X3-C motif) receptor 1	1524	601470
CXCR6	Chemokine (C-X-C motif) receptor 6	10663	605163
DAG1	Dystroglycan 1 (dystrophinassociated glycoprotein 1)	1605	128239
ENTPD	Ectonucleoside triphosphate diphosphohydrolase 3	956	603161
GNAT1	Guanine nucleotide binding protein, alpha transducing activity polypeptide 1	2779	139330

GPX1	Glutathione peroxidase 1	2876	138320
LAMB2	Laminin, beta 2 (laminin S)	3913	150325
MAP4	Microtubule-associated protein 4	4134	157132
PH4	PH-4 hypoxia-inducible factor prolyl 4-hydroxylase	54681	N/A
PLXNB1	Plexin B1	5364	601053
RASSF1	Ras association (RalGDS/AF-6) domain family 1	11186	605082
RHOA	Ras homolog gene family, member A	387	165390
RIS1	TMEM158 transmembrane protein 158 (RIS-1 Ras-induced senescence 1)	25907	N/A
RPL29	Ribosomal protein L29	6159	601832
RPSA	Ribosomal protein SA	3921	150370
SEMA3F	Semaphorin 3F	6405	601124
Scotin	Scotin	51246 6	07290
SEMA3B	Sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphoring) 3B	7869	601281
STAB1	Stabilin 1	23166	608560
TRAIP	TRAF interacting protein	10293	605958
VIPR1	Vasoactive intestinal peptide receptor 1	7433	192321

N/A, not applicable. GeneID, and OMIM identities are indicated.

Table 2. *Mutations identified in TREX1 in RVCL.*

#	Mutation	Frameshift	Reference	Geographical (Background)
1	3688_3689insG	V235	Grand et al	North America (European)
2	3688_3689insG	V235	Storimans et al Terwindt et al	Netherlands
3	3727_3730dupGTCA	T249	Jen et al	North America (Chinese)
4	3688_3689insG	V235	Grand et al	North America (Ashkenazi-Jewish)
5	3691_3692insA	T236	Weil et al	Germany
6	3835_3836insA	R284	Cohn et al	Australia
7	3688_3689insG	V235	Unpublished	North America
8	3688_3689insG	V235	Unpublished	Australia
9	3843_3844insG	L287	Unpublished	Netherlands

Supplementary Videos 1-4 Legends.

Supplementary Video 1

Confocal microscopy video showing functional consequences of RVCL associated *TREX1* mutations as modeled in transiently expressed HEK293T cells. Fluorescent protein (green) and TOPRO3 stained nuclei (blue). Fluorescence expression pattern of fluorescent protein (FP) alone. The protein is diffusely distributed in the cytoplasm and in the nucleus.

Supplementary Video 2

Confocal microscopy video showing functional consequences of RVCL associated *TREX1* mutations as modeled in transiently expressed HEK293T cells. Fluorescent protein (green) and TOPRO3 stained nuclei (blue). Fluorescence expression pattern of wild type TREX1 tagged with the fluorescent protein (FP-TREX1). This fusion protein is found in a perinuclear compartment and is excluded from the nucleus.

Supplementary Video 3

Confocal microscopy video showing functional consequences of RVCL associated *TREX1* mutations as modeled in transiently expressed HEK293T cells. Fluorescent protein (green) and TOPRO3 stained nuclei (blue). Fluorescence expression pattern of mutant TREX1, tagged with the fluorescent protein (FP-V235 fs). The mutant form of TREX1 exhibits an expression pattern identical to the fluorescent protein (FP) alone.

Supplementary Video 4

Confocal microscopy video showing functional consequences of RVCL associated *TREX1* mutations as modeled in transiently expressed HEK293T cells. Fluorescent protein (green) and TOPRO3 stained nuclei (blue). Fluorescence expression pattern of carboxylterminal 106 amino acids of TREX1 tagged with the fluorescent protein (FP-C-106). This fusion protein is found in a perinuclear compartment and is excluded from the nucleus. This pattern is identical to the native exonuclease, implicating this short stretch of amino acids in mediating the perinuclear localization of the protein.

Supplemental Methods

Samples

We analyzed DNA samples from nine families with clinical symptoms of RVCL (Supplementary Table 2). Informed consent was obtained from all patients, in accordance with procedures and regulations of the Institutional Review Boards.

Mutation detection

Genomic DNA was isolated from peripheral blood leukocytes or immortalized cell lines from consenting subjects as approved by IRB. In specific cases (Washington University Genome Sequencing Center), Phi29-based whole genome amplification was performed on genomic DNA samples. Gene sequences of candidate genes were extracted from GenBank (www.ncbi.nlm.nih.gov) and Ensembl (www.ensembl.org) databases. Primers to amplify the coding exons and exon-intron boundaries of candidate genes were designed with the PrimerDesign script that is based on the use of the Primer 3 program. Universal (forward and reverse) tails were added to the 5′ ends of amplification primers to serve as the sequencing primer sites (primer sequences and PCR conditions are available on request). Direct sequencing of purified PCR products was done by using dye-terminator chemistry and electrophoresed on a MegaBase500 (Amersham Biosciences, Princeton, NJ) capillary sequencer or either the ABI3700 or ABI3730 automated sequencer (Applied Biosystems, Foster City, CA). The sequence traces were assembled and scanned for variations from the reference sequence using the PolyScan informatics suite or Vector NTI suite 9.0.0 program (Invitrogen, Carlsbad, CA). The tagged variations were then manually reviewed. All detailed protocols are available on request.

For further mutational analysis of *TREX1*, primers were designed to amplify the coding exons of *TREX1* (Supplementary Table 3). Purified PCR amplification products were sequenced using dye-terminator chemistry and electrophoresed on a MegaBase500 (Amersham Biosciences, Princeton, NJ) capillary sequencer or an ABI3700 sequencer (Applied Biosystems, Foster City, CA). Sequencing was analyzed using PolyPhred. Anonymized control samples were screened by sequencing. Controls were matched with the ethnic origins of the mutations. The RVCL *TREX1* mutations were not detected in 192 Caucasian (HD100CAU, Coriell), 192 Chinese (HD100CHI, Coriell) and 300 Dutch control alleles.

Supplementary Table 3. *Primers and PCR conditions for TREX1 exon.*

Primer	Primer Sequence Forward	Primer Sequencen Reverse
1	tgtaaaacgacggccagtatggtggtgagagggacagacc	caggaaacagctatgaccaaagatgagggtctgcatggg
2	tgtaaaacgacggccagtgaatgtgctggtcccactaagg	caggaaacagctatgaccaaggctaggagcaggttggc
3	tgtaaaacgacggccagtctctccctgtgtgtggctcc	caggaaacagctatgaccttgtgacagcagatggtcttgg
4	tgtaaaacgacggccagtctaggcagcatctacactcgcc	caggaaacagctatgaccatcctgctagggaaagtgaggg

The appropriate universal sequencing primer was used for either reading the forward or reverse strand of all amplicons. Forward sequencing primer (5′-TGTAAAACGACGGCCAGT-3′); reverse sequencing primer (5′-CAGGAAACAGCTATGACC-3′). Incubation conditions: temperature, 60°C; Mg^{2+} concentration, 1.5 mM.

Haplotype analysis

For 20 microsatellite markers in the chromosome 3p21.1-p21.3 region, standard PCRs were performed using a PTC200 thermal cycler (Bio-Rad Laboratories, Foster City, CA). PCR products were analyzed on an ABI3700 sequencer (Applied Biosystems, Foster City, CA) and genotypes were assigned using GENESCAN and GENOTYPER software (Applied Biosystems, Foster City, CA). Two investigators scored genotypes independently. In addition, 13 SNPs in the region closely flanking *TREX1* were typed by direct sequencing. Disease haplotypes were constructed by inspection of segregation.

Molecular cloning, mutagenesis, expression, and purification of E. coli proteins

Mutations (V235fs, T249fs and R114H) were constructed using the cDNA clone encoding TREX1 variant 1 (Origene TC304415) as a template for site-directed mutagenesis. Oligos used for mutagenesis were as follows:

V235fs (5' CATGTATGGGGGTCACAGCCTCTG 3' and 5' CAGAGCGTGTGACCCCCATACATG 3')

T249fs (5' TCTGCTGTCAGTCACAACCACTGC 3' and 5' GCAGTGGTTGTGACTGACAGCAGATG 3')

R114H (5' AGCCTTCCTGCGGCACCAGCCACAGCCCTGG and 3' ACCAGGGCTGTGGCTGGTGCCGCAGGAAGGC).

In the case of TREX1, V235fs, T249fs and R114H, the inserts were subcloned by PCR using the TREX1 cDNA clone as a template into the E. *coli* expression vector pET28a+-1 {a derivative of pET28a+ (Novagen) created in house} containing an Nterminal 6x His epitope tag. Correct clones were transformed into the E. *coli* strain BL21CodonPlus (DE3)-RIL (Stratagene, La Jolla, CA). Cells containing the TREX1 plasmids were grown at 37°C to an $Absorbance_{600}$ of 0.6. Isopropyl-1-thio-β-D4 galactopyranoside was added to a final concentration of 1 mM and the cultures incubated at 37°C for an additional 3h. Cells were then harvested and the pellets were stored at -80°C. For purification of recombinant proteins, cells were resuspended in cold sonication buffer (50 mM Tris pH 8.0, 500 mM NaCl, 10% glycerol, 5 mM beta mercaptoethanol, 1 mM imidazole, and 1mM PMSF). The cell suspension was sonicated and centrifuged at 15,000 g at 4°C for 20 min to obtain a cleared lysate. Histagged proteins were batch adsorbed to Ni-NTA Agarose (Qiagen, Valencia, CA) for 1 h at 4°C. The beads were washed extensively in wash buffer (sonication buffer containing 25 mM imidazole) and packed into a 5 ml polypropylene (Qiagen) column. After additional washes, fractions were collected during elution with five column volumes of elution buffer (sonication buffer containing 250 mM imidazole). Fractions containing His-TREX1 proteins were identified by SDS-PAG electrophoresis and Western blotting with an anti-HIS-HRP conjugated antibody (Clontech, Mountain View, CA) or, alternatively, with Coomassie Brilliant Blue staining. The samples were pooled, concentrated and aliquots of the purified proteins frozen at -80°C.

Exonuclease assays

1 µg of Poly(dA) (GE Healthcare, Princeton, NJ) was labeled at the 3' end with ^{32}P dATP (GE Healthcare) using Terminal Transferase (Roche Diagnostic Corp., Indianapolis, IN). Reactions containing 50 mM Tris pH 8.5, 4 mM $MgCl_2$, 1 mM DTT, 10 µg BSA, 0.01 g radiolabeled poly(dA) substrate, and recombinant exonuclease TREX1 protein were incubated in a total volume of 100 l at 37°C. Aliquots were removed at the indicated times and ethanol precipitated in the presence of 50 µg denatured calf thymus DNA (Sigma-Aldrich, St. Louis, MO). Ethanol-soluble radioactivity released into the supernatant was measured by scintillation counting.

Generation of N-terminally-tagged TREX1 constructs

To directly visualize TREX1 within the living cell, all TREX1 forms were N-terminally tagged with the enhanced yellow fluorescent protein (eYFP). For clarity, the epitope tag is designated hereafter as fluorescent protein (FP) tag. At the wavelength employed it gives green fluorescence. The FP coding sequence was excised via EcoRI/BsrGI from sT-DAF-eY [1]. This fragment was utilized in a three-fragment ligation with EcoRI/XbaI-digested CD59dGPI [2] and the respective *Bsr*GI/*Xba*I-digested PCRderived TREX1 forms (see below). This resulted in aminoterminal tagging of TREX1 FP-TREX1, FP-V235fs and FP-T249fs). Wild-type FP was expressed from the second cistron of sT-DAF. Wild-type TREX1 (FP-TREX1) was used as a template DNA to generate the Aicardi-Goutieres R114H mutant (FP-R114H) by site-directed mutagenesis as described above. To study the effect on cellular localization of the carboxyl-terminus of TREX1, a FP-tagged fusion protein containing the last 106 amino acids of native TREX1 was generated (FPC106). A 646 bp *Bsr*GI/*Bsa*I fragment was excised from FP-TREX1, to remove the entire aminoterminus including all exonuclease sites. The cohesive ends were then blunted and the linearized 4532 bp fragment relegated, resulting in FP-C106. All PCR-derived DNA fragments and ligation products were verified by DNA sequencing.

Confocal Microscopy

HEK293T cells (7 x 10^5 on cover slides in 6-well plates) were transiently transfected overnight with 1.5 µg of each construct using TransIT-293 (Mirus, Madison, WI), according to manufacturer's directions. Following two washes in PBS, the cells were fixed for 30 min at room temperature (RT) in PBS containing 2% paraformaldehyde. Following two washes in PBS, the cover slides were incubated for 30 min at RT in PBS containing a 1/2000 dilution of Topro3 (Molecular Probes, Carlsbad, CA) to visualize the nuclei, washed again with PBS and mounted on slides over night with ProLong Gold (Molecular Probes). Samples were examined using a Zeiss LSM 510 laser scanning confocal microscope and images were processed using Image Examiner Software (Zeiss, Jena, Germany).

Expression, SDS-PAGE, and Western blotting

Following overnight transient transfection of HEK293T cells (described above), cells were washed with PBS, lysed with 1% Nonidet P-40, 0.05% SDS in PBS with 2 mM PMSF for 15 min at 4°C, and centrifuged at 12,000 g for 10 min. Supernatants were immediately evaluated or frozen at -80 °C. The Western blot was loaded with 5 x 10^5 cell equivalents per lane on non-reduced and electrophoresed (10% SDS-PAG). Following transfer to nitrocellulose, the blots were probed with 1:4000 dilution of monoclonal anti-GFP antibody JL-8 that recognizes both GFP and YFP (Clontech Laboratories) and then HRP donkey anti-mouse IgG (Amersham Biosciences).

References

1. Spitzer D et al (2004) *Molecular Immunology* 40, 911-919.

2. Spitzer D, Hauser H & Wirth D (1999) *Human Gene Therapy* 10, 1893.

Accession codes

GenBank: cDNA and amino acid numbering was determined using the TREX1 protein AAK07616 and nucleotide sequence NM_033627 (with the A at 2986 as the first base of the initiating ATG codon).

URLs

The UCSC Genome Browser is available at http://genome.ucsc.edu/.

The Marshfield chromosome 3 genetic map is found at http://research.marshfieldclinic.org/genetics.