ORIGINAL ARTICLE

Exome sequencing identifies mutations in the gene *TTC7A* in French-Canadian cases with hereditary multiple intestinal atresia

Mark E Samuels,[1] Jacek Majewski,[2] Najmeh Alirezaie,[2] Isabel Fernandez,[1,3] Ferran Casals,[1] Natalie Patey,[1,4] Hélène Decaluwe,[1,5] Isabelle Gosselin,[6] Elie Haddad,[1,3,5] Alan Hodgkinson,[1] Youssef Idaghdour,[1] Valerie Marchand,[1,5] Jacques L Michaud,[1,5] Marc-André Rodrigue,[6,7] Sylvie Desjardins,[6] Stéphane Dubois,[6] Francoise Le Deist,[1,3] Philip Awadalla,[1,5] Vincent Raymond,[6,7] Bruno Maranda[8]

▶ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/jmedgenet-2012-101483).

[1]Centre de Recherche du CHU Ste-Justine, University of Montreal, Montreal, Quebec, Canada
[2]Department of Human Genetics, McGill University, Montreal, Quebec, Canada
[3]Department of Microbiology and Immunology, University of Montreal, Montreal, Quebec, Canada
[4]Department of Pathology, University of Montreal, Montreal, Quebec, Canada
[5]Department of Pediatrics, University of Montreal, Montreal, Quebec, Canada
[6]Department of Neurosciences, Centre de recherche du CHU de Québec, Université Laval, Québec City, Quebec, Canada
[7]Département de Médecine Moléculaire, Université Laval, Québec City, Quebec, Canada
[8]Medical Genetics Service, University of Sherbrooke, Sherbrooke, Quebec, Canada

Correspondence to
Bruno Maranda,
Medical Genetics Service, University of Sherbrooke, Sherbrooke, Quebec, Canada J1K 2R1; Bruno.Maranda@USherbrooke.ca

Received 14 December 2012
Accepted 22 January 2013
Published Online First 19 February 2013

To cite: Samuels ME, Majewski J, Alirezaie N, *et al*. *J Med Genet* 2013;**50**:324–329.

ABSTRACT

Background Congenital multiple intestinal atresia (MIA) is a severe, fatal neonatal disorder, involving the occurrence of obstructions in the small and large intestines ultimately leading to organ failure. Surgical interventions are palliative but do not provide long-term survival. Severe immunodeficiency may be associated with the phenotype. A genetic basis for MIA is likely. We had previously ascertained a cohort of patients of French-Canadian origin, most of whom were deceased as infants or in utero. The goal of the study was to identify the molecular basis for the disease in the patients of this cohort.

Methods We performed whole exome sequencing on samples from five patients of four families. Validation of mutations and familial segregation was performed using standard Sanger sequencing in these and three additional families with deceased cases. Exon skipping was assessed by reverse transcription-PCR and Sanger sequencing.

Results Five patients from four different families were each homozygous for a four base intronic deletion in the gene *TTC7A*, immediately adjacent to a consensus GT splice donor site. The deletion was demonstrated to have deleterious effects on splicing causing the skipping of the attendant upstream coding exon, thereby leading to a predicted severe protein truncation. Parents were heterozygous carriers of the deletion in these families and in two additional families segregating affected cases. In a seventh family, an affected case was compound heterozygous for the same 4bp deletion and a second missense mutation p.L823P, also predicted as pathogenic. No other sequenced genes possessed deleterious variants explanatory for all patients in the cohort. Neither mutation was seen in a large set of control chromosomes.

Conclusions Based on our genetic results, *TTC7A* is the likely causal gene for MIA.

INTRODUCTION

Hereditary multiple intestinal atresia (MIA) (MIM (243150)) is a severe congenital disorder, first formally defined in 1973.[1] MIA typically involves multiple lesions which occur at various levels throughout the small and large intestines.[2] Although surgical intervention is sometimes attempted, outcomes are poor and the condition is usual fatal within the first month of life. To date, no primary aetiology has been proved for the condition. Importantly, in some cases MIA is also associated with either mild or severe combined immunodeficiency (SCID),[2–5] raising the possibility that an abnormal immune response might be the origin of the disease. However failure of normal embryogenesis remains a possibility with inflammation or immune dysfunction only secondary. Most MIA cases are sporadic, although a genetic basis for congenital cases is likely given some familial recurrences.[2 6] A cluster of cases from Quebec, ascertained by us, is also consistent with a genetic cause, given frequent documented French-Canadian founder effects going back to the first immigrants and leading to increased incidences of specific recessive disorders.[7] Other neonatal conditions involving local atresias within the small or large intestine are also known, but it remains unclear whether these represent fundamentally different disease entities from MIA or simply less severe examples.

Through whole exome sequencing of patients of French-Canadian origin, we identified two mutations in a single gene, the tetratricopeptide repeat (TPR) domain 7A gene *TTC7A*, that could explain the disease in the affected cases. Mouse mutations in the orthologous gene are known and while the mice exhibit some abnormalities in their immune system, they lack a gastrointestinal phenotype.[8–10]

METHODS

Clinical ascertainment and patient descriptions

In the course of clinical practice over many years, more than 20 patients with MIA have been identified at the Centre Hospitalier de l'Université Laval (CHUL) in Québec City, some previously described.[2] All cases were reportedly of French-Canadian origin, though not formally related by known genealogy up to their great-grandparents. DNA suitable for genome-wide analysis was obtained from these cases as well as three newly ascertained cases. DNA from three additional deceased cases was obtained from paraffin blocks (F2, F4 and F5). We also had intestinal tissues embedded in plastic for two deceased cases (F3, F5) but DNA extracted from these blocks was

degraded and could not be amplified. DNA from the sampled patient in F2 was also degraded. DNA for segregation studies was obtained for parents and/or unaffected siblings from all seven families (F1 to F7).

The three cases which underwent exome sequencing are described. The first case was born at term. The baby was immediately brought to medical attention because of intestinal obstruction and omphalocoele. He died the following day of bowel obstruction. Autopsy revealed MIA. The second case is a fetus aborted at 23 weeks of gestation. Fetal ultrasound showed significant bowel distension at 22 weeks. Since the parents had a child who died at 2 days of MIA the previous year, a recurrence was highly suspected. Fetal autopsy showed MIA. The third case was born after a 35 weeks pregnancy, characterised by polyhydramnios. Intestinal obstruction was suspected at birth on clinical grounds. Imaging findings were consistent with a diagnosis of MIA. Considering the invariably fatal outcome of this condition in our centre, palliative care was given to the infant. Autopsy was declined by the parents and the baby died at home during the 1st week.

Patients with MIA have also been identified at the Centre Hospitalier Ste-Justine (CHU Ste-Justine) in Montreal.[1] In one family (F7) two newly ascertained male siblings (P11 and P14), not previously described, were diagnosed with MIA. DNA was obtained from peripheral blood of one sibling, and from preserved tissue of the other. The family is reportedly of French-Canadian origin on the maternal side, though not known to be related to the CHUL and partial English ancestry on the paternal side.

The first sibling was born after 32 weeks of pregnancy complicated by polyhydramnios. Prenatal diagnosis of intestinal atresia was highly suspected because bowel distension was observed on fetal ultrasound at 25 weeks of gestation. Surgery performed at 2 days of life consisted in resection of multiple atresia involving the jejunum and the colon, and an ileostomy was created. Two subsequent surgeries were performed at days 16 and 35 to treat new atresias involving the pylorus and the jejunum. After resection, the patient presented bowel obstruction and died at day 47. The second sibling was born at 34 weeks after delivery was induced because of fetal distress. MIA was suspected based on echography performed at 22 weeks of pregnancy. During surgery performed at 1 day of life, multiple atresias involving the pylorus, ileum, proximal colon and rectum were resected. The progression of the disease required additional surgery at day 28. In spite of the treatment with steroids and cyclosporin A introduced at the age of 40 days because of a severe inflammation observed in biopsies, the intestinal disease progressed. The patient also had a severe immunodeficiency characterised by recurrent infections associated with hypogammaglobulinaemia and profound T cell lymphopenia. Haematopoietic stem cell transplantation was performed using a human leukocyte antigen (HLA) 6/6 matched cord blood at 6.5 months. The intestinal disease progressed and the patient died at 1 year of age.

Molecular studies

Informed written consent was obtained for genomic molecular studies using peripheral blood, immortalised cell lines or fixed tissue samples from deceased cases, and for studies of living first degree relatives. Genomic DNA was extracted from peripheral blood using the Puregene DNA isolation protocol. DNA was extracted from intestinal tissues embedded in paraffin or plastic using several rounds of phenol extraction and ethanol.

Exon regions were captured using the Agilent SureSelect All Exon 50 MB (V3) exome enrichment kit and sequencing of 100 bp paired end reads was carried out on Illumina HiSeq 2000. Capture and sequencing was performed at the McGill University and Genome Quebec Centre for Innovation. Sequence reads were aligned to the human reference genome hg19 using Burrows-Wheeler Alignment (BWA 0.5.9).[11] Genome Analysis Toolkit (GATK) was used for realignment around suspected indels,[12] and Picard was employed to mark duplicate reads and exclude them from downstream analyses.[13] Single nucleotide variants (SNVs) and short insertions and deletions (indels) were called using Samtools mpileup.[13] Mutations were annotated using Annovar[14] and custom scripts. Filtering of common variants and recurrent technical artefacts of sequencing chemistry and/or read misalignment was performed using dbSNP132, Exome Variant Server (EVS), the 1000 Genomes variant dataset and a custom database of approximately 700 exomes previously sequenced at our centre for unrelated projects. Variants seen in more than 20 control exomes, SNVs with a ratio of variant reads less than 0.2 and insertion/deletions with a ratio of variant reads less than 0.15, were excluded from further analysis. Only variants predicted to change protein sequence (exonic non-synonymous SNVs, short indels, splice site SNVs) were considered. Private variants were defined as those absent from dbSNP, 1000 Genomes, Exome Variant Server and our internal exome databases. Variants were visualised using Integrative Genomics Viewer (IGV), or with NextGene (Soft Genetics, Inc.)

Sequences of primers for amplification and sequencing of *TTC7A* exons are given in online supplementary materials. Genomic DNA was extracted using the Puregene DNA isolation protocol from 28 ml of whole blood drawn by venipuncture. PCRs were performed on a Hybaid Omnigene Temperature Cycling System in a total volume of 50 ml containing 100 ng of genomic DNA, 20 pmol of each primer, 200 mM dNTPs, 50 mM KCl, 10 mM Tris (pH 9), 1.5 mM MgCl2, 0.01% gelatin, 0.1% Triton X-100 and 1 U Taq polymerase (Invitrogen, Burlington, ON). Amplifications were carried out using a 'hot-start' procedure. Taq polymerase was added after a 5 min denaturation step at 95°C. Samples were then processed through 35 cycles of denaturation (95°C for 30 s) and annealing (55°–60°C for 30 s), followed by one last step of elongation (72°C for 50 s). PCR products were diluted in five volumes of phosphate buffered (PB) buffer (Qiagen, Mississauga ON), transferred on a Whattman GF/C filter plate, washed twice with 80% ethanol/20 mM Tris (pH 7.5) and eluted in 50 ml of water. Samples were quantified by the PicoGreen reagent protocol. A second PCR was performed on Applied Biosystems Gene Amp PCR System 9700 (96 wells) or 9700 Viper (384 wells) machines to incorporate the sequencing dyes using a protocol of 25 cycles of denaturation (95°C for 10 s) and annealing (55°C for 5 s), followed by one last step of elongation (59°C for 2 min). PCR products were purified by the Applied Biosystems ethanol-EDTA precipitation protocol, collected using a Beckman-Coulter Allegra 6R centrifuge, and resuspended in a 50% HiDi-formamide solution. Samples were then run on Applied Biosystems Prism 3700xt DNA Analyzer automated sequencers. Sequence data were analysed using the Staden preGap4 and Gap4 programmes, or else using MutationSurveyor (Soft Genetics, Inc.)

Amniocytes were cultured from the exome-sequenced affected patient of pedigree 4 using standard culture procedures. For splicing analysis, RNA was extracted from cells and reverse transcription-PCR (RT-PCR) was performed using the Qiagen OneStep RT-PCR kit (Qiagen). The RT-PCR forward primer was designed in exon 4 with a few nucleotides extending in exon 5 to span the splice junction. Reverse primer was designed in exon 12

with a few nucleotides extending in exon 11 to span the splice junction (see online supplementary materials). This strategy allowed the screening of alternative splicing events in encompassed exons, and minimised amplification of genomic DNA.

Multiple sequence alignments were performed using ClustalOmega[15] or Multiple Sequence Comparison by Log Expectation (MUSCLE),[16] together with BoxShade. Putative *TTC7A* orthologs were obtained from the National Center for Biotechnology Information (NCBI) Gene and Homologene databases for the following species: *Homo sapiens* (modern human), *Pan troglodytes* (common chimpanzee), *Macaca mulatta* (rhesus monkey), *Bos taurus* (domestic cow), *Canis lupus familiaris* (domestic dog), *Felis cattus* (domestic cat), *Equus caballus* (horse), *Loxodonta africana* (African elephant), *Oryctolagus cuniculus* (European rabbit), *Otolemur garnettii* (northern greater galago), *Mus musculus* (house mouse), *Rattus norvegicus* (brown rat), *Ailuropoda melanoleuca* (giant panda), *Monodelphis domestica* (grey short-tailed opossum), *Sarcophilus harrisii* (Tasmanian devil), *Meleagris gallopavo* (wild turkey), *Anolis carolinensis* (green anole lizard), *Danio rerio* (zebrafish), *Oreochromis niloticus* (Nile tilapia). The putative ortholog of *Gallus gallus* (common chicken or red jungle fowl) is incomplete due to a problem with genomic annotation of the critical carboxy terminal coding exon and was omitted from alignments.

Missense mutation pathogenicity prediction was performed using PolyPhen2,[17] Protein Variation Effect Analyzer (PROVEAN),[18] MutationTaster,[19] Sorting Intolerant from Tolerant (SIFT)[20 21] and PMUT.[22] Splice site mutation pathogenicity prediction was performed using MaxEntScan,[23 24] and with the Berkeley neural network algorithm NNSplice, V.0.9.[25] Because one putative mutation is intronic, DNA sequence numbering was based on annotated genomic contig NC_000002.11. Protein sequence numbering was based on RefSeq entry NP_065191.2.

RESULTS

We first performed genome-wide homozygosity mapping using 408 microsatellite markers spaced at an average of 10 cm with DNA from three affected patients. As no obvious region was detected, we then sequenced the exomes of these patients from different families of the CHUL (F1, F4, F6, see figure 1). An average of 12.5 gigabases of raw sequence were acquired for each sample. After aligning reads to the consensus human genome using an implementation of the Burrowss-Wheeler algorithm (BWA), average mean read depth of bases in consensus coding sequence exons assessed by GATK was 91X. Average consensus coding sequence bases covered by at least 5, 10 and 20 reads were 94%, 91% and 86, respectively.

As in other published studies, each individual exome sample contained more than 100 000 called variants. After filtering with public and custom exome variant databases to remove common SNVs and recurrent technical artefacts, each sample contained about 400 rare non-synonymous variants (predicted to alter protein function or splicing). Only one variant was shared in the homozygous state among all three patient exomes, this was a four base pair deletion c.53344_53347delAAGT in the gene *TTC7A*, encoding TPR protein 7A (MIM[609332], figure 1, see online supplementary figures S1, S3A). No gene contained two different pathogenic variants (potential compound heterozygotes) in all three patient exomes. The 4 bp deletion occurs at the 5′ splice donor site of exon 7, where the consensus genome sequence contains a directly repeated AAGT forming the GT splice donor site. The deleted allele contains

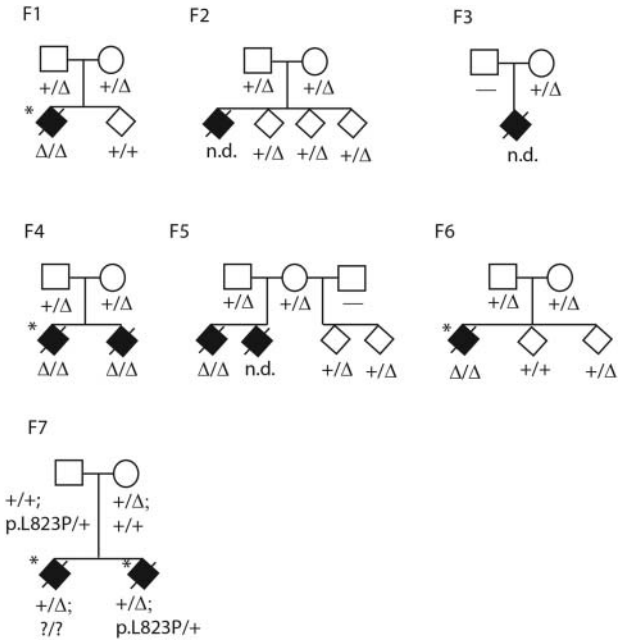


Figure 1 Multiple intestinal atresia pedigrees F1–F7. F1–F6 are from Centre Hospitalier de l'Université Laval cohort, F7 is from Centre Hospitalier Ste-Justine cohort. Affecteds are shown as filled diamonds, status of 4 bp deletion in exon 7 only is shown in F1–F6 as a delta sign, status of the 4 bp exon 7 deletion and p.L823P missense in exon 20 are shown in F7. Horizontal dash or question mark means no genotype was obtained. Probands who underwent exome sequencing are starred. One affected was previously reported (Bilodeau *et al*, case 11).[2] To preserve confidentiality sex is not shown, and children are not in obligate order of birth.

only one AAGT. The variant falls in a region of extended homozygosity of several megabases in each of the three patients based on exome-derived high quality variant calls. The exact length of homozygosity varies among the three patients, while the minimal overlapping shared region of homozygosity was approximately 1.2 Mb. For the most part all three affecteds share minor SNV alleles identical by state across this region, although exome software is not optimised for such an analysis.

To validate the homozygous AAGT deletion, all 20 coding exons of the *TTC7A* gene and their intron-exon junctions were screened for variations by PCR-based Sanger sequencing in the three exome sequenced samples and available members of their nuclear families (F1, F4, F6). The only variation detected in all three patients was the 4 bp homozygous deletion. The other 19 coding exons had no interesting sequence variants. In all three families, both parents were heterozygous for the deletion and unaffected siblings were either heterozygous for the deletion or homozygous wild type (figure 1). In family F2, DNA of the affected deceased case was unavailable but both parents were heterozygous for the deletion and unaffected siblings were heterozygous for the deletion (figure 1). In family F3, one available parent was heterozygous for the deletion. DNA from two additional MIA patients (families F4 sibling of exome sequenced patient, F5 one of two affected patients) were extracted from embedded intestinal tissue biopsies. Both these patients also carried the homozygous AAGT deletion, and parents of the F5 patient were both heterozygous for the deletion.

Although the deleted allele retains a potential GT splice donor site, the mutation is predicted to have a deleterious effect on splicing, with a maximum entropy score of −9.7 according

to MaxEntScan. Similarly, the Berkeley splice site prediction programme NNSplice correctly identified the splice donor site at the distal end of exon 7 using 132 bp of wild type genomic sequence bracketing the site, but did not predict any donor site in the same region with the 4 bp deletion. To assess potential aberrant splicing, poly-A^+ RNA was extracted from Hela cells or from cultured amniocytic cells of one patient (from pedigree 4) and from one normal control infant. Following RT-PCR, *TTC7A* RNA was detected from the patient cells, as well as from control and Hela cells. Consistent with the bioinformatics predictions, a smaller size product was observed from RNA of the patient, versus RNA from control or Hela cells, for PCR reactions of cDNA spanning exon 7 (see online supplementary figure S2). Sequencing of the smaller product verified that exon 7 was skipped, with splicing of exon 6 directly to exon 8 (see online supplementary figure S3B). This aberrant splicing generated a 158 bp deletion in the resulting cDNA, predicted to cause a frameshift, 281 amino acids through the gene with 49 new amino acids followed by a termination codon, with a presumptive severe effect on protein function (full length wild type sequence is 858 amino acids). In human lymphocyte RNAseq data from other unrelated projects, we examined the state of *TTC7A* splicing and found rare examples of reads skipping exon 7 (four reads excluding exon 7 vs 880 reads including the exon), suggesting that exon 7 inclusion may be intrinsically variable even with a wild type splice donor site.

An additional family (F7) with two MIA affected siblings was independently ascertained at CHU Ste-Justine. Exome sequencing identified the same 4 bp AAGT deletion seen homozygous in the CHUL, but in the heterozygous state in both siblings. The entire *TTC7A* gene could not be interpreted due to variable exon sequence coverage. All 20 *TTC7A* coding exons and their intron-exon junctions were therefore screened for variations by Sanger sequencing, confirming that the two siblings were both heterozygous for the 4 bp AAGT deletion (figure 1, see online supplementary figure S4). A second heterozygous mutation was found in both siblings: c.A133074G; p.L823P, in carboxy-terminal coding exon 20 of the gene (figure 1, see online supplementary figure S3C). By Sanger sequencing, each parent was heterozygous for one of the two mutations, the 4 bp deletion on the maternal side consistent with her French-Canadian origin, and the missense on the paternal side. Together the results are consistent with the mutations being compound heterozygous in the two affected patients of this family, although we were unable to test the exon 20 mutation in the other affected sibling directly. The leucine at position 823 in the human gene sequence is highly conserved among putative TTC7A orthologs (figure 2), and mutation p.L823P is predicted as probably damaging by PolyPhen2 HumDiv and HumVar, as deleterious by PROVEAN (score −5.7, anything lower than −2.5 is considered deleterious), as damaging by SIFT, as disease causing by MutationTaster and as pathological by PMUT. In the *TTC7A* gene itself, SNPs identified as homozygous in the three homozygous affecteds were heterozygous in the two affecteds of the compound heterozygotes, or else had low overall read coverage.

Neither the exon 7 AAGT deletion nor the exon 20 missense mutation in *TTC7A* is found in dbSNP V. 135 (which includes data from the 1000 Genomes project), and neither was observed in 1500 French-Canadian control chromosomes (although as noted the paternal missense mutation may be of British or Scottish origin).[26]

Based on our molecular genetic results, *TTC7A* is a strong candidate for the MIA phenotype in all our analysed cases.

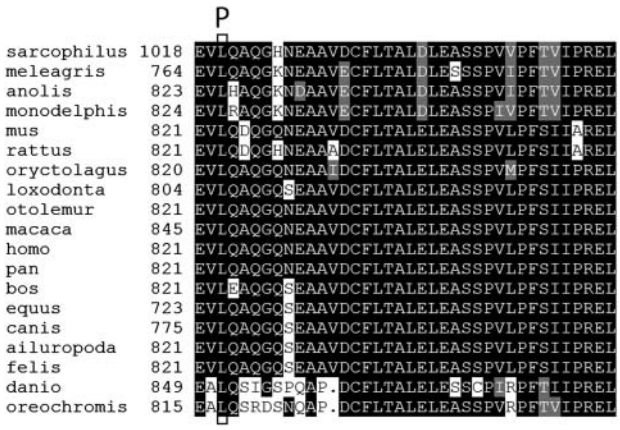


Figure 2 Multiple sequence alignment of putative *TTC7A* orthologs, showing the C-terminus of the protein. Residue L823 is boxed in all orthologs, and the mutation p.L823P (according to the human amino acid numbering) is shown above the top row.

DISCUSSION

Exome and PCR-based sequencing studies of seven cases of MIA, as well as obligate carriers and unaffected relatives from seven families, all from French-Canadian Quebec, identified two different mutations segregating in the gene encoding TTC7A. One mutation, a 4 bp deletion of one copy of a duplicated tetranucleotide encompassing a splice donor site, was predicted to be pathogenic by two different splice prediction tools. Five of the cases were homozygous for this mutation, which was also found in the heterozygous state all available parents including those of additional affecteds whose DNA was not directly available. The variant was not seen in the homozygous state in any unaffected family members of probands. This deletion has also not been seen in any other exome sequence data in our laboratories, which include a large number of French-Canadian samples. In direct RT-PCR experiments, the aberrant splicing product was observed in cultured cells from one of the patients, consistent with the bioinformatics predictions. Skipping exon 7 leads to a predicted severe protein truncation. Although the aberrantly spliced product should be susceptible to nonsense-mediated decay, sufficient RNA was nonetheless present in the patient's cells to permit the assay. The second mutation, missense p.L823P, occurs at a highly conserved residue near the carboxy terminus of the protein, and is predicted to be pathogenic by five different programmes. Two cases, siblings in the same family, were likely compound heterozygotes for the 4 bp intronic splice site deletion and the missense mutation (second hit verified in one sibling). These results, together with SNP calls derived from the exome data (not extensively revalidated), are consistent with *TTC7A* mutations being causal for the intestinal atresia in all of these cases. It is likely that the 4 bp deletion represents a founder mutation in the French-Canadian population, with the missense mutation having arisen more recently or else derived from a British or Scottish population; similar observations of combined founder and non-founder mutations in the same gene have been observed for other monogenic phenotypes in eastern Canada.[27] It remains to be seen whether all ascertained cases of congenital MIA are caused by mutations in *TTC7A* or whether additional genetic or non-genetic aetiologies are also significant.

Little is known concerning the function of *TTC7A* in human development. The TPR domain consists of approximately 34

amino acids with a consensus sequence of (WLF)-X(2)-(LIM)-(GAS)-X(2)-(YLF)-X(8)-(ASE)-X(3)-(FYL)-X(2)-(ASL)-X(4)-(PKE) according to the NCBI conserved domain database.[28] This appears to be a protein-protein interaction domain with various potential partners involved in different cellular processes including cell cycle control, phosphate turnover and subcellular localisation,[29 30] thus it does not suggest any particular physiological role for *TTC7A* itself. Multiple TPR domains occur at the carboxy-terminal end of TTC7A, and the missense putative mutation p.L823P is in one of these. Physical studies indicate that TPR domains possess significant α-helical structure,[29] hence since proline residues are strong helix-breakers the L-to-P mutation is likely to disrupt the structure of that TPR repeat. TTC7A appears widely expressed at low levels, though at slightly higher levels in myeloid and monocyte cells according to the Bio Gene Portal System (BioGPS) human tissue expression database.[31]

Interestingly, three spontaneously arising mutations in the mouse *TTC7A* ortholog, *Ttc7*, are known. The flaky skin (*fsn*) recessive mouse was originally described in 1995, with a phenotype including a papulosquamous disease of the skin, anaemia with an associated haematopoietic disorder and gastric forestomach hyperplasia.[10] Subsequently another spontaneous recessive mouse mutant, hereditary erythroblastic anaemia (*hea*) was found and determined to be allelic to *fsn*.[9] The *hea* strain shares the haematopoietic and skin aspects of the phenotype,[32] but not reportedly the stomach hyperplasia,[9] although the skin phenotype was only observed when the mutation was transferred from the original CFO genetic background to C57Bl/6J.[9] The common causal gene for *fsn* and *hea* was positionally cloned as *Ttc7*, the mouse ortholog of human *TTC7A*.[33] The *hea* allele results from a large deletion encompassing exons 1–14, whereas the *fsn* allele results from an intronic transposon insertion that leads to incorporation of a transposon-derived extra exon between mouse coding exons 14 and 15; this leads to an inframe insertion of 61 amino acids into the protein disrupting one of the TPR domain repeats.[33 34] In terms of phenotypic pleiotropy the *fsn* allele seems slightly more severe, although the *hea* mice reportedly have shorter life spans. The *fsn* mice provide an important model system for human psoriasis, and it is unclear whether the anaemia and associated iron deficiency is a primary or secondary aspect of the mouse mutations. The *fsn* mice have abnormal lymphoid cell populations including a reduced ratio of CD4 to CD8 cells and massively enlarged peripheral lymph nodes, moreover lymphoid cells are hyper-responsive to interleukin growth factors.[35] A third allelic mouse mutation in the gene was later identified, *int*, and shown to result from deletion of exons 9–10, leading to an inframe protein deletion.[8 36] The *int* mice have severe haematopoietic and papulosquamous skin disorders and die as neonates, however they do not appear to exhibit a gastrointestinal phenotype.

These mouse mutants exhibit quite different clinical presentations from human patients. The mouse genetics do indicate that *Ttc7* plays a role in proper immune system function, which is consistent with an immune involvement in MIA. However, the immune phenotype of the mice is different from what is observed in human patients, who present with a SCID-like phenotype in most of the cases in which an immunodeficiency is observed. In addition, the mice have no documented intestinal involvement (although the *fsn* strain does have forestomach papillomas), whereas the human patients have no obvious skin involvement. Conceivably these differences are due to allele-specific effects on immune function; perhaps the human mutations are hypomorphic compared to the mouse mutations. The lack of intestinal involvement in the mice especially in the large deletion strain is difficult to reconcile, although the much shorter intestinal tract and differences in development of intestinal mucosa and immune defence in rodents may be relevant. Alternatively, even the severe mouse mutations show significant genetic background effects, thus modifiers elsewhere in the genome may be relevant to the human phenotype of MIA. Treatment of *fsn* mice with helminth (worm)-derived glycans ameliorates the skin phenotype.[37] Helminth extracts have been tested in humans for other autoimmune disorders with some promising results.[38] Given the typically fatal course of MIA and lack of any therapeutic regime other than pre-emptive surgery to remove damaged tissue, it is worth speculating whether glycan treatments either prenatally or postnatally might be of benefit. Finally, although stem cell transplantation during the 1st year of life for the SCID was unsuccessful in one patient, conceivably fetal stem cell transplantation through the fetal cord might represent a therapeutic option, if indeed the intestinal phenotype is secondary to an immune system dysfunction. This however remains to be established.

Acknowledgements Foremost we thank the families who generously contributed their time and materials to this research study. Thanks to Maryse Lagace, Lysanne Patry, Renée Dicaire and Christine Massicotte for assistance with patient and sample coordination and ethics submission. The FORGE Canada Consortium Steering Committee consists of Kym Boycott (leader; Ottawa, Ontario), Jan Friedman (co-leader; Vancouver, British Columbia), Jacques Michaud (co-leader; Montreal, Quebec), Francois Bernier (Calgary, Alberta), Michael Brudno (Toronto, Ontario), Bridget Fernandez (St Johns, Newfoundland), Bartha Knoppers (Montreal, Quebec), Mark Samuels (Montreal, Quebec), Steve Scherer (Toronto, Ontario). We also thank Janet Marcadier (Clinical Coordinator) and Chandree Beaulieu (Project Manager) for their contribution to the infrastructure of the FORGE Canada Consortium.

Contributors AH participated in bioinformatic analyses of whole exome data, BM was responsible for clinical ascertainment and phenotype definition, and participated in manuscript preparation, EH participated in clinical ascertainment and phenotype definition and in manuscript preparation, FC participated in exome sequencing, molecular studies and data analysis, FLD was responsible for clinical ascertainment and phenotype definition and participated in manuscript preparation, HD participated in clinical ascertainment and phenotyping patients, IF participated in phenotype definition and manuscript preparation, IG assisted with clinical ascertainment, phenotype definition, patient sampling and ethics approvals, JM assisted in clinical ascertainment and manuscript development and participated in manuscript preparation, JLM coordinated bioinformatic analysis of whole exome sequence data and participated in manuscript preparation, MAR participated in bioinformatic analyses of whole exome data and in molecular studies, MES coordinated molecular studies and primary manuscript development, NA participated in bioinformatic analyses of whole exome data, NP participated in phenotyping patients, VR supervised molecular validation and RNA-based studies and participated in manuscript preparation, PA coordinated bioinformatic analysis of whole exome sequence data and participated in manuscript preparation, SyD participated in molecular studies, StD participated in molecular studies, VM participated in clinical ascertainment and phenotyping patients, YI participated in exome sequencing, molecular studies and data analysis.

Funding This work was funded by the Government of Canada through the Canadian Institutes of Health Research, Genome Canada, Genome Quebec, Genome British Columbia and the Ontario Genomics Institute (OGI-049). MES was supported by the Centre de Recherche du CHU Ste-Justine. PA, FLD and EH were supported by CIHR Grant 120299. EH is a scholar of the Fonds de recherche du Quebec-Sante. VR was supported by Fonds de recherche du Québec- Santé.

Competing interests None.

Patient consent Obtained.

Ethics approval This study was approved by the research ethics committees of CHU Ste-Justine and CHUL.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Data from this study may be shared with other researchers.

REFERENCES

1 Guttman FM, Braun P, Garance PH, Blanchard H, Collin PP, Dallaire L, Desjardins JG, Perreault G. Multiple atresias and a new syndrome of hereditary multiple atresias involving the gastrointestinal tract from stomach to rectum. *J Pediatr Surg* 1973;8:633–40.
2 Bilodeau A, Prasil P, Cloutier R, Laframboise R, Meguerditchian AN, Roy G, Leclerc S, Peloquin J. Hereditary multiple intestinal atresia: thirty years later. *J Pediatr Surg* 2004;39:726–30.
3 Moreno LA, Gottrand F, Turck D, Manouvrier-Hanu S, Mazingue F, Morisot C, Le Deist F, Ricour C, Nihoul-Fekete C, Debeugny P, Griscelli C, Farriaux J-P. Severe combined immunodeficiency syndrome associated with autosomal recessive familial multiple gastrointestinal atresias: study of a family. *Am J Med Genet* 1990;37:143–6.
4 Ali YA, Rahman S, Bhat V, Al Thani S, Ismail A, Bassiouny I. Hereditary multiple intestinal atresia (HMIA) with severe combined immunodeficiency (SCID): a case report of two siblings and review of the literature on MIA, HMIA and HMIA with immunodeficiency over the last 50 years. *BMJ Case Reports* 2011.
5 Cole C, Freitas A, Clifton MS, Durham MM. Hereditary multiple intestinal atresias: 2 new cases and review of the literature. *J Pediatr Surg* 2010;45:E21–4.
6 Dallaire L, Perreault G. Hereditary multiple intestinal atresia. In: Bergsma D ed. *The clinical delineation of birth defects: urinary system and others*. Baltimore: Williams & Wilkins, 1974:259–64.
7 Laberge AM, Michaud J, Richter A, Lemyre E, Lambert M, Brais B, Mitchell GA. Population history and its impact on medical genetics in Quebec. *Clin Genet* 2005;68:287–301.
8 Takabayashi S, Katoh H. A mutant mouse with severe anemia and skin abnormalities controlled by a new allele of the flaky skin (fsn) locus. *Exp Anim* 2005;54:339–47.
9 White RA, McNulty SG, Roman S, Garg U, Wirtz E, Kohlbrecher D, Nsumu NN, Pinson D, Gaedigk R, Blackmore K, Copple A, Rasul S, Watanabe M, Shimizu K. Chromosomal localization, hematologic characterization, and iron metabolism of the hereditary erythroblastic anemia (hea) mutant mouse. *Blood* 2004;104:1511–18.
10 Beamer WG, Pelsue SC, Shultz LD, Sundberg JP, Barker JE. The flaky skin (fsn) mutation in mice: map location and description of the anemia. *Blood* 1995;86:3220–6.
11 Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009;25:1754–60.
12 McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;20:1297–303.
13 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The sequence alignment/map format and SAMtools. *Bioinformatics* 2009;25:2078–9.
14 Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010;38:e164.
15 Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 2011;7:539.
16 Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;32:1792–7.
17 Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248–9.
18 Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. *PloS one* 2012;7:e46688.
19 Schwarz JM, Rodelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods* 2010;7:575–6.
20 Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 2009;4:1073–81.
21 Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res* 2012;40: W452–7.
22 Ferrer-Costa C, Gelpi JL, Zamakola L, Parraga I, de la Cruz X, Orozco M. PMUT: a web-based tool for the annotation of pathological mutations on proteins. *Bioinformatics* 2005;21:3176–8.
23 Eng L, Coutinho G, Nahas S, Yeo G, Tanouye R, Babaei M, Dork T, Burge C, Gatti RA. Nonclassical splicing mutations in the coding and noncoding regions of the ATM Gene: maximum entropy estimates of splice junction strengths. *Hum Mutat* 2004;23:67–76.
24 Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol* 2004;11:377–94.
25 Reese MG, Eeckman FH, Kulp D, Haussler D. Improved splice site detection in Genie. *J Comput Biol* 1997;4:311–23.
26 Awadalla P, Boileau C, Payette Y, Idaghdour Y, Goulet JP, Knoppers B, Hamet P, Laberge C. Cohort profile of the CARTaGENE study: Quebec's population-based biobank for public health and personalized genomics. *Int J Epidemiol* 2012 published online October 15, 2012 doi:10.1093/ije/dys160
27 Guernsey DL, Dube MP, Jiang H, Asselin G, Blowers S, Evans S, Ferguson M, Macgillivray C, Matsuoka M, Nightingale M, Rideout A, Delatycki M, Orr A, Ludman M, Dooley J, Riddell C, Samuels ME. Novel mutations in the sacsin gene in ataxia patients from Maritime Canada. *J Neurol Sci* 2010;288:79–87.
28 Marchler-Bauer A, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Liebert CA, Liu C, Lu F, Lu S, Marchler GH, Mullokandov M, Song JS, Tasneem A, Thanki N, Yamashita RA, Zhang D, Zhang N, Bryant SH. CDD: specific functional annotation with the Conserved Domain Database. *Nucleic Acids Res* 2009;37:D205–10.
29 Blatch GL, Lassle M. The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays* 1999;21:932–9.
30 Lamb JR, Tugendreich S, Hieter P. Tetratrico peptide repeat interactions: to TPR or not to TPR? *Trends Biochem Sci* 1995;20:257–9.
31 Wu C, Orozco C, Boyer J, Leglise M, Goodale J, Batalov S, Hodge CL, Haase J, Janes J, Huss JW III, Su AI. BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol* 2009;10:R130.
32 Kasahara Y, Shimizu K, Kuribayashi K. Developmental abnormalities of the thymus in hea/hea mutant mice. *Exp Anim* 2008;57:85–94.
33 White RA, McNulty SG, Nsumu NN, Boydston LA, Brewer BP, Shimizu K. Positional cloning of the Ttc7 gene required for normal iron homeostasis and mutated in hea and fsn anemia mice. *Genomics* 2005;85:330–7.
34 Helms C, Pelsue S, Cao L, Lamb E, Loffredo B, Taillon-Miller P, Herrin B, Burzenski LM, Gott B, Lyons BL, Keppler D, Shultz LD, Bowcock AM. The Tetratricopeptide repeat domain 7 gene is mutated in flaky skin mice: a model for psoriasis, autoimmunity, and anemia. *Exp Biol Med (Maywood)* 2005;230:659–67.
35 Abernethy NJ, Hagan C, Tan PL, Birchall NM, Watson JD. The peripheral lymphoid compartment is disrupted in flaky skin mice. *Immunol Cell Biol* 2000;78:5–12.
36 Takabayashi S, Iwashita S, Hirashima T, Katoh H. The novel tetratricopeptide repeat domain 7 mutation, Ttc7fsn-Jic, with deletion of the TPR-2B repeat causes severe flaky skin phenotype. *Exp Biol Med (Maywood)* 2007;232:695–9.
37 Atochina O, Harn D. Prevention of psoriasis-like lesions development in fsn/fsn mice by helminth glycans. *Exp Dermatol* 2006;15:461–8.
38 Kuijk LM, van Die I. Worms to the rescue: can worm glycans protect from autoimmune diseases? *IUBMB Life* 2010;62:303–12.

Supplementary Material

Figure S1. IGV visualization of homozygous mutation c.53344_53347delAAGT from exome data. All reads show 4 bp deletion, sequence of wild type gene and exon annotation at bottom.

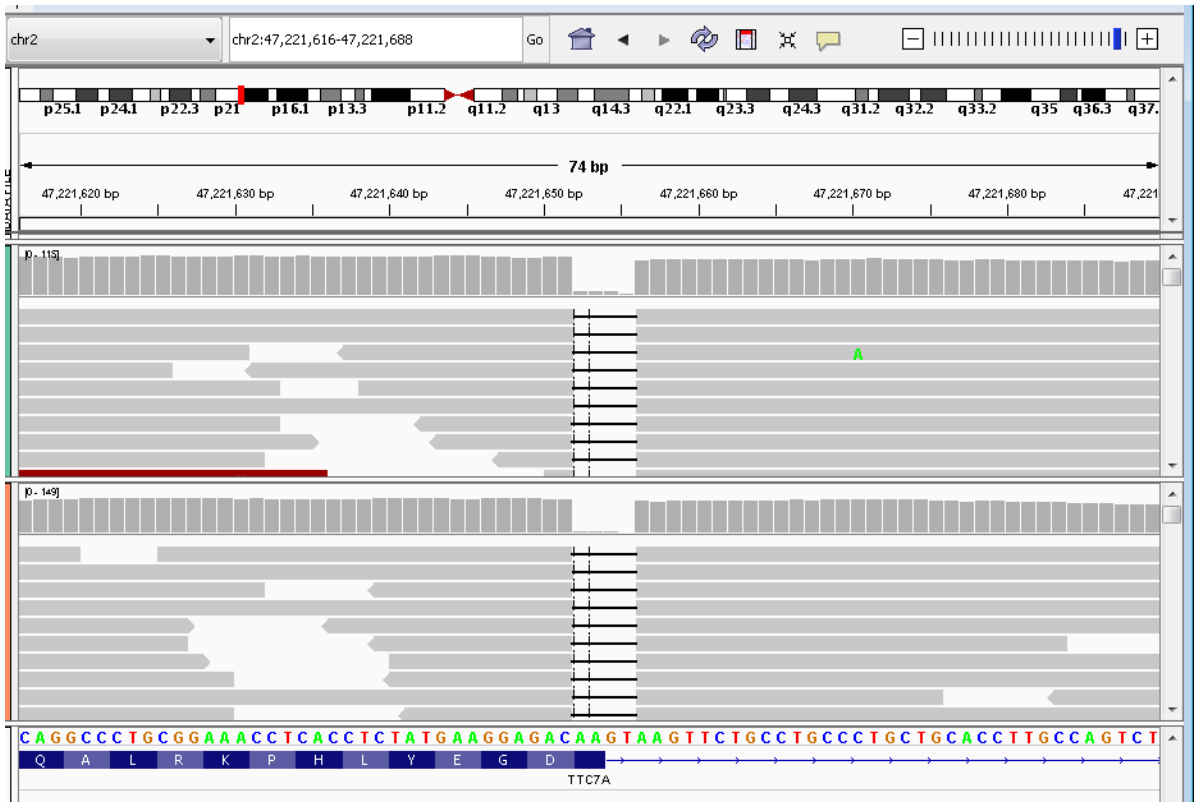


Figure S2. Agarose gel of RT-PCR reaction products from *TTC7A* encompassing exon 7. From left to right, source Hela cell RNA, 100bp size ladder (bottom band 200bp, top band 1000bp, source RNA extracted from human control amnyocytes (band is 780 bp and represents a WT spliced RNA that contains exons 4 to 12 including exon 7), source RNA of amniocytes of affected exome sequenced patient of pedigree 4 (band is 622 bp and represents an RNA that extends from exons 4 to 12 skipping exon 7 deleting 158 bp). See Materials and Methods for details on cells and RT-PCR protocol.

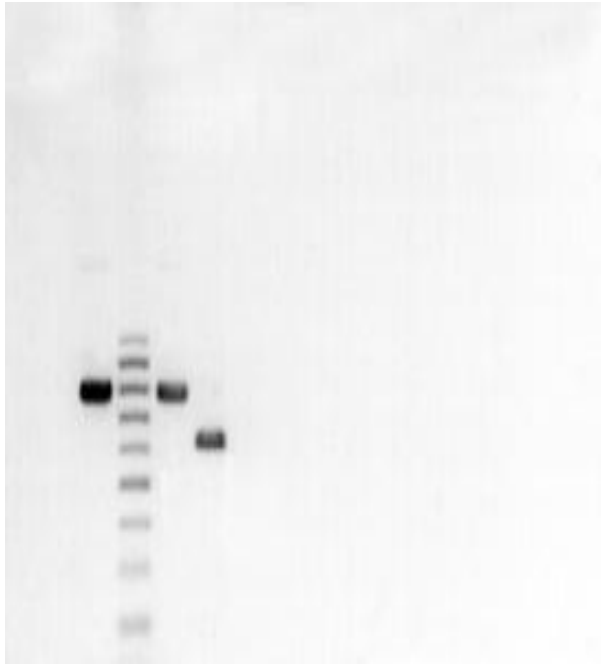
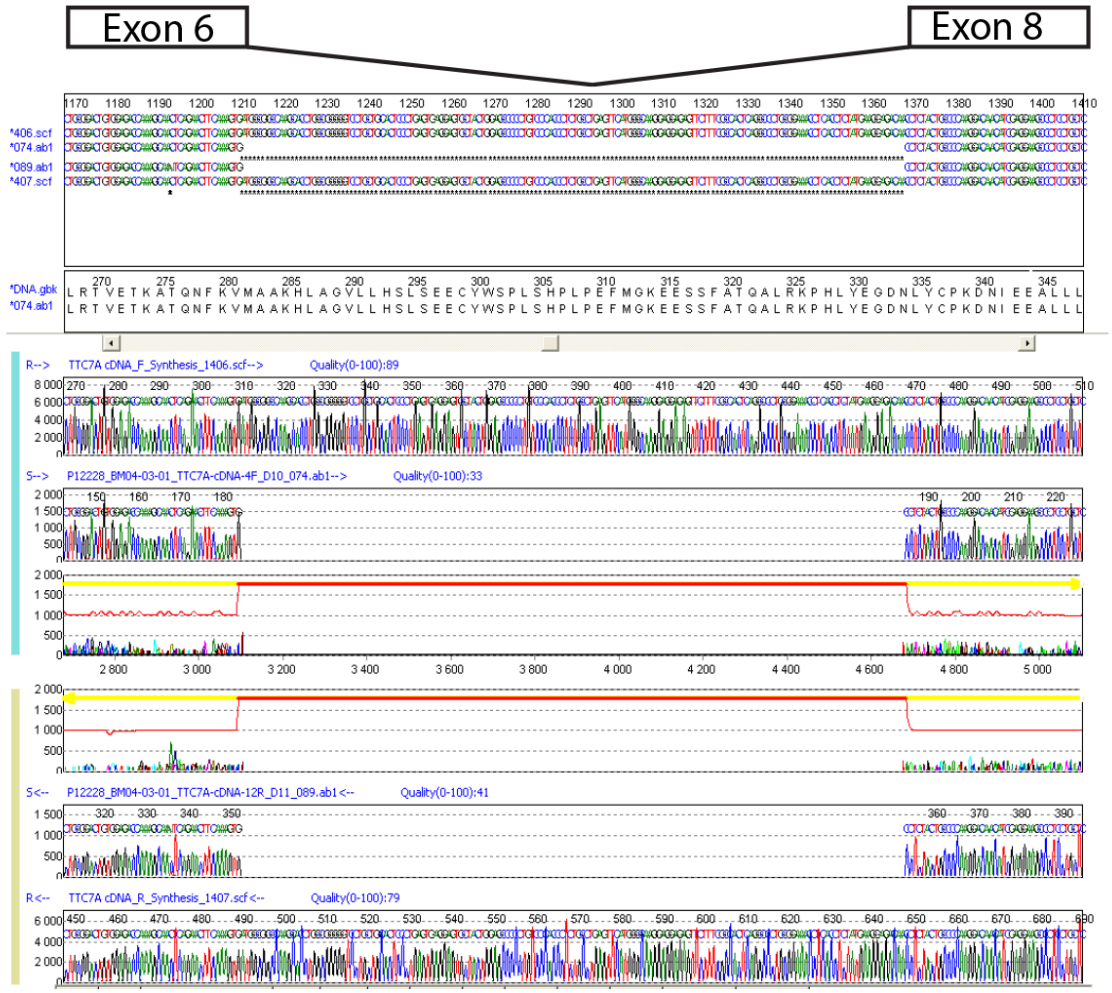


Figure S3. Sequence chromatograms of mutations in TTC7A and exon skipping products. A, homozygous mutation c.53344_53347delAAGT from sampled affected homozygous patient. B, sequence of RT-PCR products from sampled affected patient in family 4, aligned versus *TTC7A* reference cDNA to highlight exon 7 skipping. C, heterozygous mutation c.A133074G; p.L823P. Sequencing by standard PCR-Sanger fluorescent chemistry, visualized with Mutation surveyor. In each panel, top window shows nucleotide sequence of consensus from human genome project above mutant sequence, then amino acid sequence of consensus from human genome project above mutant sequence with exon number annotated. Lower window shows chromatograms for wild type sequence (virtual, generated from consensus genome project annotation), chromatogram from patient, and mutation calling by algorithm. Red arrows point to mutations in patient chromatogram.

A


*C7A.gbk
*269.scf
*018.ab1
*C7A.gbk
*018.ab1
TTC7A_R_Synthesis_53269.scf <-- Quality(0-100):91
P12042_BM01-04_TTC7A-7Rint_H04_018.ab1 <-- Quality(0-100):35

B


Exon 6
Exon 8
*406.scf
*074.ab1
*089.ab1
*407.scf
*DNA.gbk
*074.ab1
R--> TTC7A cDNA_F_Synthesis_1406.scf--> Quality(0-100):89
S--> P12228_BM04-03-01_TTC7A-cDNA-4F_D10_074.ab1--> Quality(0-100):33
S<-- P12228_BM04-03-01_TTC7A-cDNA-12R_D11_089.ab1 <-- Quality(0-100):41
R<-- TTC7A cDNA_R_Synthesis_1407.scf <-- Quality(0-100):79

C

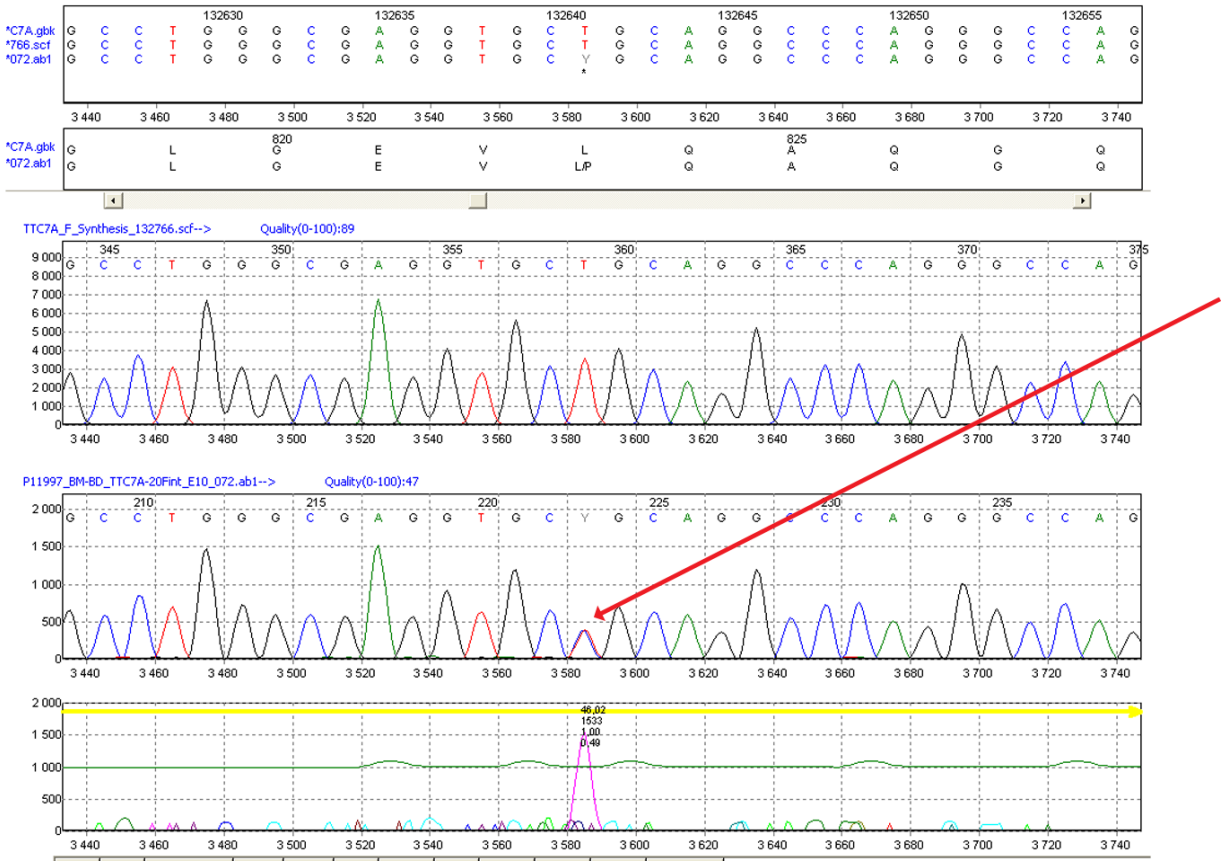

*C7A.gbk
*766.scf
*072.ab1
TTC7A_F_Synthesis_132766.scf-->
Quality(0-100):89
P11997_BM-BD_TTC7A-20Fint_E10_072.ab1-->
Quality(0-100):47

Figure S4. MutationSurveyor visualization of deconvoluted heterozygous mutation c.53344_53347delAAGT from Sanger sequencing of affected individual in pedigree 7. From top to bottom, panels are virtual wild type chromatogram, actual sequence chromatogram of heterozygous patient, deconvoluted non-deleted patient sequence, deconvoluted deleted patient sequence, deconvoluted correctly offset deleted patient sequence, mutation calling by software algorithm.

Table S1. Primer sequences for amplification and sequencing of TTC7A coding exons.

exon	PCR primers[1]		sequencing primer[1]	
1	TTC7A-exon1Fint	CTCCGCGCGGGATTAAAGT	idem	
	TTC7A-1R	GAGTCCAGCCAATGCACTGAT	idem	
			TTC7A-1Rint	TGCAGATACCACCATTTCACG
2	TTC7A-2F	CCGAGTCCTGGGAACTCTGTC	idem	
	TTC7A-2R	TGTTGCTCAGGTGACCAAATG	idem	
3	TTC7A-3F	GAGGAGGGGAGTTCTGAGCAA	idem	
	TTC7A-3R	ACCGGCATCCACCTTAGACAC	idem	
4	TTC7A-4F	CTAACACCACCGTCGCTTCAG	TTC7A-exon4Fint	TGCACCAGAGTGTCTGC
	TTC7A-4R	TGGGGGACAGAGAAGGTGACT	idem	
5	TTC7A-5F2	GCCCCTGAACAACTGGGACTGG	idem	
	TTC7A-5R2	GCACGTATGCACAGGCCAGAGA	idem	
6	TTC7A-6F	TCCTGGGCCTCTAAAGTCCTG	TTC7A-exon6Fint	CACTCCAGTTATCGAATGTGTTCA
	TTC7A-6R	TGGGAAAGGATGCTGAGATTG	idem	
7	TTC7A-7F	AGAGGGGTGGGAGCGCCAGT	idem	
	TTC7A-7R	ACCTGGACCCCAACACCTGCCT	TTC7A-exon7Rint	TCTAGTTCATCTGGCTCCTG
	TTC7A-ex7F-court	GGGTCCGAGTGCTTCCCTCT	idem	
	TTC7A-ex7R-court	GGGACAGAACCCTGCGTGAG	idem	
8	TTC7A-8F	GTGCCCATGAATTATGCAGGA	idem	
	TTC7A-8R	CCTCCAAATGAAGCCTTCGAC	idem	
9-10	TTC7A-9-10F	GGGCCTTTGTGGAGATAGGAA	TTC7A-exon10F	ATCACCCTGTGAGATCATTC
	TTC7A-9-10R	CCACAGACTCACACCAGCACA	TTC7A-exon9R	GATGGCAAAGAGCAAATG
11	TTC7A-11F	TTCTCTGCCCTGGACCTCTTC	idem	

	TTC7A-11R	CTCTCTGGAGGCCACAGTTCA	idem	
12	TTC7A-12F	GAGCACAAGGCTTCTGTTTCG	idem	
	TTC7A-12R	AGTCTGGGGGAAACAGTGGAA	idem	
13-14	TTC7A-13-14F	CTGCAGAGCTGTGTGCTTTGA	idem	
	TTC7A-13-14R	TGGAAGGCGAGGACACAAATA	idem	
15	TTC7A-15F	GGGCTCCCTTGAGTTGAGTGT	idem	
	TTC7A-15R	CCTGTCCTCACCCCTGTCTCT	idem	
16	TTC7A-16F	GCGAGTTAGGGAGGTGAGCAT	idem	
	TTC7A-16R	CGCTCTCATTCAAAAGCCTCA	idem	
17	TTC7A-17F	AGTGAGGCTGTCCCATTCTCC	TTC7A-exon17Fint	GTGAGGCTGTCCCATTC
	TTC7A-17R	AAGATCAGCCCCAGAAGTTGC	idem	
18	TTC7A-18F	CAGCTGTGGGTGAGAGGACAT	idem	
	TTC7A-18R	GCCAGAGGCGAGGACTCTGTA	idem	
19	TTC7A-19F	TGTGTCCCATGCTGTGATTTG	idem	
	TTC7A-19R	CCAGGGAACAGCAGTAGCAAG	idem	
20	TTC7A-20F	AGAGCTCCTGCAGTGGGTTTC	TTC7A-exon20Fint	AAGGCACAGTCACTTAACAG
	TTC7A-20R	CTGTGCTGAGAGTGGGGAGTG	TTC7A-exon20Rint	CATGGGTGAGGGTGAAG

1 All sequences are indicated in the 5' to 3' direction

Table S2. Primer sequences for reverse transcription, amplification and sequencing of TTC7A RNA encompassing exon 7.

TTC7A_cDNAex3F	aggcttttgtcatcaaaggcc	(mainly in exon 3 with few nucleotides overlapping the junction of exon 3-4)
TTC7A_cDNAex4F	ctgcaggaattggagaagacc	(mainly in exon 4 with few nucleotides overlapping the junction of exon 4-5)
TTC7A_cDNAex9R	ccacatctcgagttgccatg	(mainly in exon 9 with few nucleotides overlapping the junction of exon 8-9)
TTC7A_cDNAex12R	gtgctctgcttcctctagcc	(mainly in exon12 with few nucleotides overlapping the junction of exon 11-12)