

Defective complex I assembly due to *C20orf7* mutations as a new cause of Leigh syndrome

M Gerards,^{1,2} W Sluiter,³ B J C van den Bosch,^{1,2} L E A de Wit,³ C M H Calis,¹ M Frentzen,⁴ H Akbari,⁴ K Schoonderwoerd,⁵ H R Scholte,^{3,6} R J Jongbloed,¹ A T M Hendrickx,¹ I F M de Coo,⁷ H J M Smeets^{1,2}

► Supplementary tables 1 and 2 are published online only. To view these files please visit the journal online (<http://jmg.bmj.com>).

¹Department of Genetics and Cell Biology, Unit Clinical Genomics, Maastricht University, Maastricht, The Netherlands

²School for Oncology and Developmental Biology, Maastricht University, Maastricht, The Netherlands

³Department of Biochemistry, Mitochondrial Research Unit, Erasmus MC, Rotterdam, The Netherlands

⁴Unit of Botany, Institute for Biology I, RWTH Aachen University, Aachen, Germany

⁵Department of Clinical Genetics, Erasmus MC, Rotterdam, The Netherlands

⁶Department of Neuroscience, Erasmus MC, Rotterdam, The Netherlands

⁷Department of Neurology, Erasmus MC, Rotterdam, The Netherlands

Correspondence to

Dr H J M Smeets, Department of Genetics and Cell Biology, PO Box 616, 6200 MD Maastricht, the Netherlands; Bert.Smeets@molcelb.unimaas.nl

Received 5 March 2009

Revised 19 May 2009

Accepted 25 May 2009

Published Online First

18 June 2009



This paper is freely available online under the BMJ Journals unlocked scheme, see <http://jmg.bmj.com/site/about/unlocked.xhtml>

ABSTRACT

Background Leigh syndrome is an early onset, progressive, neurodegenerative disorder with developmental and motor skills regression. Characteristic magnetic resonance imaging abnormalities consist of focal bilateral lesions in the basal ganglia and/or the brainstem. The main cause is a deficiency in oxidative phosphorylation due to mutations in an mtDNA or nuclear oxidative phosphorylation gene.

Methods and results A consanguineous Moroccan family with Leigh syndrome comprise 11 children, three of which are affected. Marker analysis revealed a homozygous region of 11.5 Mb on chromosome 20, containing 111 genes. Eight possible mitochondrial candidate genes were sequenced. Patients were homozygous for an unclassified variant (p.P193L) in the cardiolipin synthase gene (*CRLS1*). As this variant was present in 20% of a Moroccan control population and enzyme activity was only reduced to 50%, this could not explain the rare clinical phenotype in our family. Patients were also homozygous for an amino acid substitution (p.L159F) in *C20orf7*, a new complex I assembly factor. Parents were heterozygous and unaffected sibs heterozygous or homozygous wild type. The mutation affects the predicted S-adenosylmethionine (SAM) dependent methyltransferase domain of *C20orf7*, possibly involved in methylation of *NDUFB3* during the assembly process. Blue native gel electrophoresis showed an altered complex I assembly with only 30–40% of mature complex I present in patients and 70–90% in carriers.

Conclusions A new cause of Leigh syndrome can be a defect in early complex I assembly due to *C20orf7* mutations.

INTRODUCTION

Leigh syndrome (MIM 256000) is an early onset, progressive, neurodegenerative disorder with a characteristic neuropathology consisting of focal, bilateral lesions in one or more areas of the central nervous system, including basal ganglia, thalamus, cerebellum, brainstem and spinal cord. The lesions are areas of demyelination, gliosis, necrosis, spongiosis, or capillary proliferation. Clinical symptoms depend on which areas of the central nervous system are affected. The most common underlying cause is a defect in oxidative phosphorylation (OXPHOS).¹ Leigh syndrome can be associated with a deficiency of any of the mitochondrial respiratory chain complexes.² The disorder is genetically heterogeneous and can be caused by mutations in the mtDNA and a variety of nuclear genes of the OXPHOS and pyruvate dehydrogenase system.

In structural complex I genes, mutations have been found in the mtDNA *ND1*,³ *ND2*,⁴ *ND3*,⁵ *ND5*,⁶ and *ND6*⁷ genes and in the nuclear *NDUFV1*,⁸ *NDUFS1*,⁹ *NDUFS3*,¹⁰ *NDUFS4*,¹¹ *NDUFS7*,¹² *NDUFS8*,¹³ *NDUFA11*¹⁴ and *NDUFA2*¹⁵ genes. In addition, mutations have been reported in the genes for the flavoprotein subunit A of complex II (*SDHA*)¹⁶ and for the complex III assembly factor *BCSL1*.¹⁷ Mutations in complex IV genes include mtDNA encoded *MTCO3*¹⁸ and nuclear encoded *COX10*,¹⁹ *COX15*,²⁰ *SCO2*,²¹ and *SURF1*,²² which are involved in complex IV assembly. Several mutations were detected in the mtDNA encoded *ATPase 6* (complex V).²³ Additional mtDNA mutations have been described for the tRNA genes for valine,²⁴ lysine,²⁵ tryptophan²⁶ and leucine.²⁷

Leigh syndrome may also be caused by mutations in components of the pyruvate dehydrogenase complex (*PDHA1* pyruvate dehydrogenase, $\alpha 1$ -subunit and *DLD*, dihydrolipoamide dehydrogenase)^{28 29} and in the gene encoding the leucine-rich PPR motif containing protein (*LRPPRC*).³⁰ Due to this extreme genetic heterogeneity, in many cases it is difficult to establish a genetic diagnosis in Leigh syndrome, especially if no additional clinical or biochemical data are available to pinpoint specific candidate genes. In order to identify the genetic cause in a consanguineous Moroccan family we performed homozygosity mapping and positional candidate gene analysis. A pathogenic amino acid substitution was identified in a new complex I assembly factor *C20orf7*³¹ which caused Leigh syndrome with a diminished complex I activity, making defective complex I assembly an important pathogenic cause of Leigh syndrome.

SUBJECTS AND METHODS

Subjects

In a consanguineous family from Morocco with 11 children, three boys presented with Leigh syndrome (figure 1), two of whom were clinically investigated. The third affected sib succumbed for unknown reason at the age of 36 years in Morocco. The disease course is comparable in the two living patients (age 29 and 23 years). At the age of 3 they presented with a progressive spasticity with involvement of arms and legs. At the age of 5 they developed a diminished facial expression. Computed tomography (CT) scanning of the brain of patient IV7 at the age of 5 years showed hypodensity and slight atrophy of the caudate nuclei and the putamen (figure 2a). There is a widening of the frontal horns of the cerebral ventricles. One year

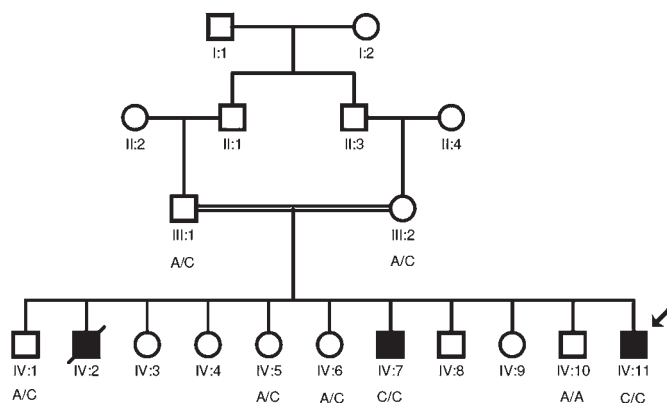


Figure 1 Pedigree of a consanguineous Moroccan family with three patients with Leigh syndrome. Black symbols indicate affected, white symbols unaffected subjects. The genotypes for the c.477A → C mutation in exon 5 of the *C20orf7* gene are indicated below the tested individuals. The proband is indicated by an arrow.

later an extrapyramidal movement disorder with dystonic posturing of the right foot was noted. In the years following, the dystonia evolved in a more choreadystonic movement disorder. A delay in mental development became apparent. The dystonia hampered swallowing, necessitating nasal tube feeding. No signs of sensory nerve conduction loss were observed. At the age of 18 years both patients were moderately retarded with an IQ of about 50, had a remarkable good temper, and were severely handicapped by the dysarthria, the dystonic posturing, and the spastic tetraplegia. Their length was just below the third centile, but this was not different from the unaffected sibs and parents.

In both patients the electroencephalogram (EEG) was normal. Pattern visual evoked responses demonstrated normal responses with N135 at 141 ms. Cranial magnetic resonance imaging (MRI) of patient IV11 at the age of 23 years showed a small residual nucleus caudatus and lesions in the basal ganglia consisting of prolongation of both T1 and T2 weighted signals in the caudate nucleus, putamen, the substantia nigra and a discrete abnormality in the peri-aqueductal grey area, and also discrete bifrontal global atrophy (figure 2b). After reaching puberty, the progression of the disease seemed to stop or slowed down. Orthopaedic interventions were necessary, because of complications of the movement disorders such as development of an equinovarus contracture in both feet and the development of a c-shaped scoliosis. Oral daily drug treatment consisted of carnitine 330 mg, riboflavin 30 mg, biotin 10 mg, and thiamine 100 mg. Parents and the other children are in good health and further family history is unremarkable. Informed consent was provided by the family for scientific investigations and publication.

Metabolic and enzymatic measurements

Routine metabolic workup was performed on the blood, urine and cerebrospinal fluid of the patients to rule out other inborn errors of metabolism including blood glucose, lactate, amino acids, acid/base status, ammonia, creatine kinase, carnitine, acylcarnitines, very long chain fatty acids, uric acid, B12, folate, cholesterol, isoelectric focusing of transferrin (for CDG, congenital disorders of glycosylation), biotinidase, lysosomal enzymes, purine and pyrimidine values; urine amino acids, organic acids, oligosaccharides, and mucopolysaccharide screening; cerebrospinal fluid (CSF) glucose, lactate, amino acids, and neurotransmitter measurements. A needle muscle biopsy from vastus lateralis muscle was performed at 22 and 29 years of age in patients IV11 and IV7, respectively. Complex I activity was determined in

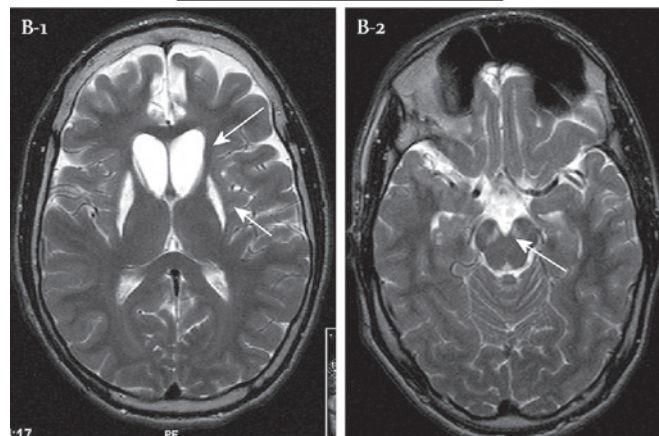
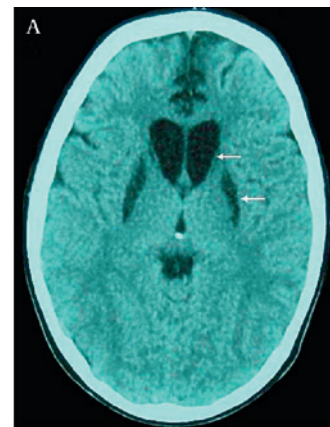


Figure 2 Computed tomography (CT) and magnetic resonance imaging (MRI) of patients IV7 and IV11, respectively. (A) CT scan of patients IV7 brain at the age of 5 years. The CT scan shows a hypodensity and slight atrophy of the caudate nuclei and the putamen (arrows) and a widening of the frontal horns of the cerebral ventricles. (B-1) Cranial MRI of patient IV11 at the age of 23 years. Prolongation of T2 weighted signals in the residual part of the nucleus caudatus and putamen and (B-2) at the level of the midbrain of the substantia nigra (arrow).

duplicate in these biopsy specimens and in peripheral blood lymphocytes (PBMCs) of patients IV7 and IV11 and their unaffected brother IV10. Assays to determine complex I and citrate synthase activities and protein content of the mitochondrial fractions were performed as described before.^{32–35}

Fibroblasts were grown in Dulbecco's modified Eagle's medium (DMEM; Gibco by Invitrogen, Carlsbad, CA, USA) supplemented with 10% fetal bovine serum (FBS, Gibco), 50 µg/ml of uridine (Acros Organics Geel, Belgium) and 50 units/ml penicillin and 50 µg/ml streptomycin (BioWhittaker, Walkersville, MD, USA). To isolate mitochondria, cells were resuspended in isolation buffer containing 0.25 M sucrose, 10 mM Tris-HCl pH 7.5, and 1 mM EDTA and homogenised on ice (10 strokes at 1500 rpm). Homogenates were centrifuged at 1600 g at 4°C for 10 min to remove cell debris and nuclei. Subsequently, mitochondria were pelleted from the supernatant at 10 000 g at 4°C for 10 min. Cardiolipin synthase (CLS) activity was determined in mitochondrial membranes after swelling the isolated organelles twice in 10 mM Bis-Tris propane-HCl buffer pH 7.4 and sedimenting mitochondrial membranes, which were finally resuspended in this Bis-Tris propane-HCl buffer containing 50% glycerol to a protein concentration of about 1 mg protein/ml. Mitochondrial protein, 1–6 µg, was used for CLS assays in a total volume of 50 µl, as described before³⁶ except that the reaction mixtures were incubated at 37°C for 1 h.

Homozygosity mapping

Homozygosity mapping was performed, using the Affymetrix GeneChip Human Mapping 10K 2.0 Array (Santa Clara, CA, USA) for a whole genome analysis. Samples were processed and labelled according to the instructions of the manufacturer, hybridised in a GeneChip hybridisation oven followed by wash and stain with the GeneChip Fluidics Station 450, and scanning with the GeneChip Scanner 3000 (Affymetrix). Genotypes were generated by the GeneChip DNA analysis software (GDAS). The Copy Number Analysis Tool (CNAT, Affymetrix) was used to detect homozygosity regions in patient samples. Candidate regions were defined as homozygosity regions present in the patient samples but not in other family samples. Parametric lod scores were calculated using the Merlin package (version 1.1.2)⁵⁷ with a recessive disease model.

Mutation analysis

The mtDNA was screened for deletions by long range polymerase chain reaction (PCR) and heteroplasmic point mutations by denaturing high performance liquid chromatography (DHPLC) analysis as described before.³⁸ The exons and flanking introns of the human *CRLS1* and *C20orf7* genes were amplified with specific intronic primers (supplementary data table 1).

PCR was performed with 50 ng DNA using Taq-polymerase and buffer (Invitrogen, Carlsbad, CA, USA) with a final concentration of 1.5 mM MgCl₂. Cycle conditions were: 94°C for 5 min, followed by 33 cycles of 94°C for 1 min, 54°C for 1 min, and 72°C for 1.5 min with a final elongation step of 72°C for 7 min. PCR products were directly sequenced with the PRISM Ready Reaction Sequencing Kit (Applied Biosystems, Foster City, CA, USA) on an ABI3100 automatic sequencer (Applied Biosystems).

Mutation specific restriction digestion

A mutation specific restriction digestion assay was developed for the c.477A→C substitution in the *C20orf7* gene. A total of 10 µl PCR product of exon 5 was digested for 1 h at 37°C with NheI and MseI, followed by a heat inactivation for 20 min at 65°C. MseI does not cleave if the mutation is present. NheI was added to eliminate a constant band of 127 bp, which would otherwise trouble the interpretation. After digestion, the samples were resolved on a 3% agarose gel presenting a fragment of 109 bp for wild type samples compared to a 130 bp fragment for the mutant.

Blue native-polyacrylamide gel electrophoresis (BN-PAGE)

Mitochondria were isolated from PBMCs^{33,34} and incubated for 10 min on ice in lysis buffer (50 mM NaCl, 5 mM aminocaproic acid, 50 mM imidazole, pH 7.0) containing 1% Triton X-100 and supplemented with complete protease inhibitors (Roche Diagnostics GmbH, Mannheim, Germany), 1 mM 4-(2-aminoethyl)-benzenesulphonyl-fluoride hydrochloride (Roche) and 2 mM diisopropyl fluorophosphate (Fluka Chemica, Sigma-Aldrich, St. Louis, MO, USA). Next, 2.5 µg of PBMC mitochondrial protein was processed for one dimensional BN-PAGE as previously described³⁹ using NativePAGE 3–12% gels (Invitrogen, Breda, The Netherlands) and Western transferred on PVDF membrane. Immunoreactive proteins were detected by mouse monoclonal antibodies against complex I subunits ND1 (Santa Cruz Biotechnology, Santa Cruz, CA, USA), NDUFA9 and complex II subunit 70 kDa Fp (MitoSciences, Eugene, OR, USA). Proteins were visualised using biotinylated secondary antibodies against mouse IgG (Amersham Biosciences, Little Chalfont,

Bucks, UK), streptavidin-biotinylated horseradish peroxidase complex (Amersham) and SuperSignal West Femto (Pierce Biotechnology, Rockford, IL, USA). The molecular weights of the protein bands were estimated by a NativeMark Unstained Protein Standard (Invitrogen). Images were acquired using a calibrated densitometer (GS-800; Bio-Rad, Bio-Rad Laboratories, Hercules, CA, USA) and quantified using PDQuest Advanced software package (version 8.0.1 build 055; Bio-Rad).

RESULTS

Metabolic and enzymological studies of two sibs with Leigh syndrome

Routine metabolic workup in urine and blood plasma of patients IV7 and IV11 only revealed a slight increase in blood alanine values without a raise in blood or urine lactic acid. Glucose tolerance test was normal. Also the methionine, vitamin B12, folic acid, purines and pyrimidines, and S-adenosylmethionine (SAM) concentrations were in the normal range (86–128 nmol/l). Cerebrospinal fluid examination at the age of 5 years of patient IV7 showed an increased lactic acid of 5 mmol/l (controls <2.0 mmol/l). Complex I activity, normalised by citrate synthase activity, was in muscle and PBMCs of patient IV7 respectively 36% and 6% of the controls and in patient IV11 respectively 48% and 33%. Activities of other OXPHOS complexes were only available for patient IV11 and were 105%, 123% and 87% of the control values for complex II, III and IV, respectively. Morphological studies show an increase in average fibre diameter. Staining for succinate tetraxolium reductase and NADH tetraxolium reductase showed increased subsarcolemmal activities. Staining for cytochrome *c* oxidase demonstrated some COX negative fibres.

Homozygosity mapping in the Leigh syndrome family

Molecular genetic analysis was initially directed at the mtDNA. First, the entire mtDNA of the muscle sample of patient IV7 was tested for deletions by long range PCR, but no mtDNA deletion was observed (data not shown). Next, the mtDNA was screened by DHPLC analysis,³⁸ but no heteroplasmic point mutations were detected, making the involvement of a nuclear gene likely. Both parents, two affected (IV7 and IV11) and four unaffected sibs (IV1, IV5, IV6 and IV10) were tested with the Human Mapping 10K GeneChips. CNAT analysis revealed a large homozygosity region of 11.5 Mb on chromosome 20 (from 2.4 Mb to 13.9 Mb, supplementary data table 2), yielding a parametric LOD score of 2.3. This region contained a total of 111 genes. The selection of genes with a putative mitochondrial function was based on the MitoP2 database (<http://www.mitop.de:8080/mitop2/>) and the list of Calvo *et al.*⁴⁰ which is based on a variety of criteria, including mitochondrial import sequence and co-expression with known mitochondrial genes. A total of eight candidate genes—*IDH3B* (isocitrate dehydrogenase 3 [NAD⁺] β-subunit), *MRPS26* (mitochondrial ribosomal protein S26), *PANK2* (pantothenate kinase 2), *CRLS1* (cardiolipin synthase 1), *CDS2* (CDP-diacylglycerol synthase (phosphatidate cytidyltransferase) 2), *HAO1* (hydroxyacid oxidase (glycolate oxidase) 1), *PAK7* (p21 (CDKN1A)-activated kinase 7), and *C20orf7*—were selected for mutation analysis.

Mutation detection in eight candidate genes

Exons and flanking intron sequences of the eight candidate genes were analysed by conventional sequence analysis. Only two potentially pathogenic variants were found in the *CRLS1* and *C20orf7* gene, while in the other six genes no pathogenic

mutations were identified. Patient IV11 was homozygous for a c.592C → T mutation in exon 4 of the *CRLS1* gene, changing proline at position 193 into leucine. Patient IV7 was homozygous mutant as well, both parents and two sibs were heterozygous, and one sib was wild type. The mutation was neither reported before nor present in any of the single nucleotide polymorphism (SNP) databases (NCBI SNP, Genome Variation Server and MutationDiscovery.com). Analysis of a Moroccan and Dutch control population (resp. 92 and 294 alleles) using sequence analysis revealed an allele frequency of, respectively, 20% and 10%, which was unexpected given the absence of the variant in the SNP databases. CLS activity was reduced to 50–56% in fibroblasts of patients compared to unrelated wild-type control fibroblasts. Based on the high frequency of the (heterozygous) mutation in the controls and the residual activity of CLS of about 50%, we concluded that this mutation could not by itself explain the rare clinical features in this family.

A second homozygous mutation was identified in the *C20orf7* gene. Both patients (IV7 and IV11) were homozygous for an A-to-C transversion at nucleotide 477 in exon 5 of isoform 1 (NM_024120.3), whereas both parents and two other sibs were heterozygous and one sib was wild type (figures 1 and 3a). Two isoforms of *C20orf7* are currently known, consisting of, respectively, 345 amino acids (isoform 1 (NM_024120.3), containing exons 1–11), and 317 amino acids (isoform 2 (NM_001039375.1), containing exons 1–4 and 6–11). The mutation is present in isoform 1 only. The mutation changed a leucine at position 159 to phenylalanine in the highly conserved (from man to *Drosophila*) S-adenosylmethionine (SAM) dependent methyltransferase domain of the protein (figure 3b). Analysis with the program SOPM⁴¹ shows that the secondary structure is likely to change due to breakage of an α helix in the mutant protein (data not shown). SOPM uses sequence information and similarities to predict secondary protein structures. The mutation was not present in 110 Moroccan and 312 Dutch alleles, supporting a pathogenic role for the mutation.

p.L159F substitution in *C20orf7* disturbs complex I assembly

Recently, Sugiana *et al.*⁵¹ reported that *C20orf7* resides within the mitochondrial matrix with a function in the assembly or stability of an early complex I assembly intermediate that contains among others ND1. Therefore we tested if the p.L159F substitution affected complex I assembly. BN-PAGE (Blue-Native polyacrylamide gel electrophoresis) was performed in PBMCs of patients IV7 and IV11 (homozygous mutant), heterozygous carriers III1, III2 and IV6 and the homozygous wild type sib IV10 with antibodies against ND1 and NDUFA9 to detect the mature 880 kDa complex I. BN-PAGE showed a decrease of mature complex I in patient samples IV7 and IV11 to 30–40% of the control values. In carriers (III1, III2 and IV6) this was 70–90% of the normal amount of complex I (figure 4A, B).

DISCUSSION

Leigh syndrome is a progressive neurodegenerative disorder, which can be caused by mutations in the mtDNA and a variety of known and still unknown nuclear genes. Due to this genetic heterogeneity it is often difficult to establish a genetic diagnosis and characterise the pathological process involved. This study reports an early complex I assembly defect due to a missense mutation in the new complex I assembly factor *C20orf7* as the pathological cause. The c20orf7 mutation affects the highly conserved S-adenosylmethionine (SAM) dependent methyltransferase domain, which may be involved in methylation of

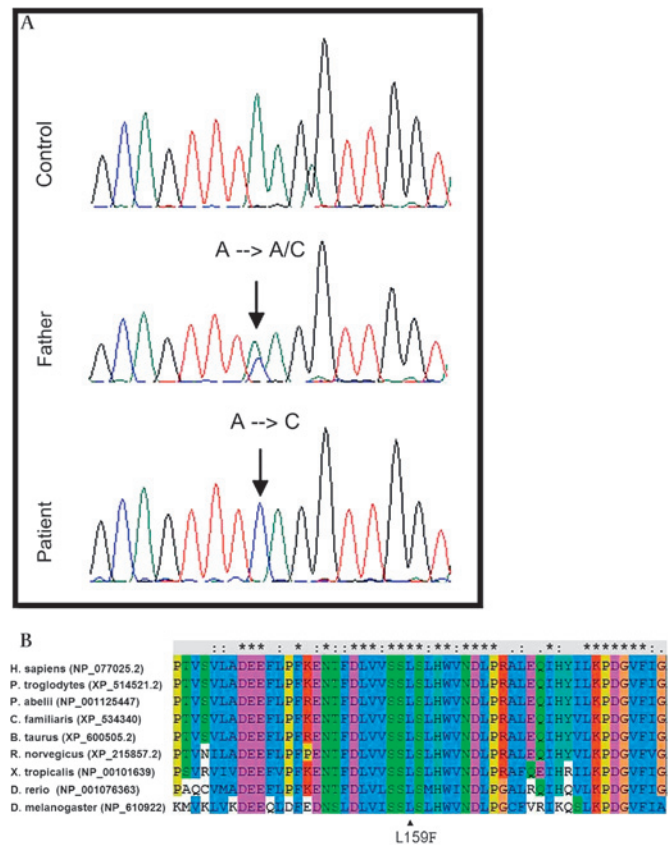


Figure 3 Mutation analysis in the *C20orf7* gene reveals a c.477A → C mutation changing leucine at position 159 to phenylalanine. (A) Sequence analysis of exon 5 of the *C20orf7* gene. (B) Conservation of leucine 159 from man to *Drosophila* (ClustalW).

NDUFB3 as a critical step in early complex I assembly as postulated by Sugiana *et al.*⁵¹ The patients do not display the full picture of Leigh syndrome and clinically the phenotype has overlap with infantile bilateral striatal necrosis (IBSN, MIM 271930), but the involvement of structures in the brain stem make it more compatible to Leigh syndrome. Defects in complex I assembly/stability leading to enzymatic deficiency have been reported before in Leigh syndrome due to mutations in structural or accessory subunits.⁴² To our knowledge this is, next to c8orf38,⁵¹ the second defect in a genuine complex I assembly factor in Leigh syndrome, which adds defective complex I assembly as an important pathological concept to the growing list of pathogenic causes of Leigh syndrome.

When these studies were in progress, a family with a lethal neonatal form of complex I deficiency with a mutation in the *C20orf7* gene was reported.⁵¹ Patients were homozygous for a missense mutation in exon 7 of the *C20orf7* gene (NM_024120.3), leading to a substitution of leucine 229 to proline, which is present in both isoforms. Patients' fibroblasts almost completely lacked the mature complex I holo-enzyme. ND1 or an intermediate complex containing ND1 was not present, suggesting an early assembly defect. In concordance with their observation, a decrease of mature complex I relative to complex II down to 30–40% compared to normal is observed in our patients' PMBCs as determined by BN-PAGE.

The clinical manifestation of the p.L159F mutation in this family is less severe than the neonatal lethal p.L229P mutation, and also the complex I deficiency and assembly defect are less pronounced. Either the mutation or the affected domain of

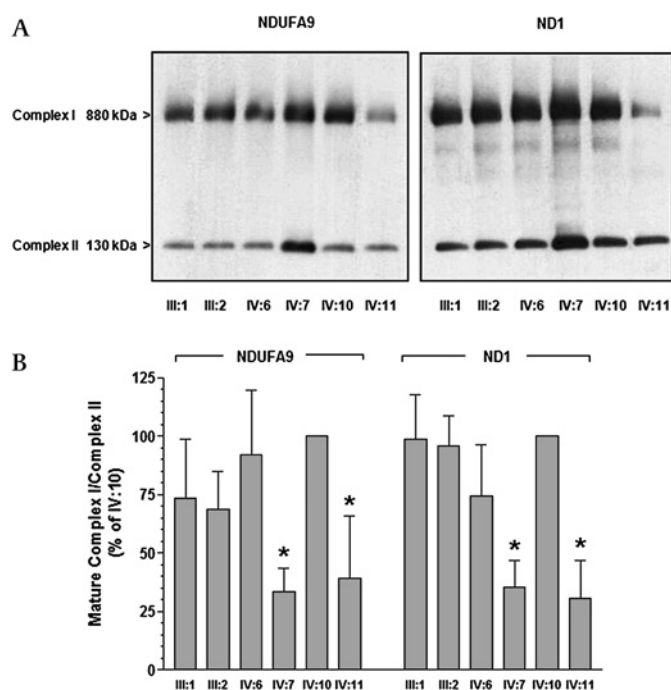


Figure 4 Complex I assembly is altered in patient PBMC. Mitochondria were isolated from PBMC of wild-type (A/A), carriers (A/C) and a patient (C/C) donors, solubilised in 1% Triton X-100, and subjected to BN-PAGE and Western immunoblot analysis and quantified. Complex I detected by an antibody to the 39 kDa subunit (NDUFA9) or ND1 subunit is at approximately 880 kDa, and complex II detected by an antibody to the 70 kDa subunit is at approximately 130 kDa. (A) After 10–30 s exposure time. (B) Histogram of A showing the density volumes of complex I normalised by complex II as percentage of the wild-type (IV:10) donor, with error bars representing the SEM of three or four gels. * $p < 0.05$ lower than the wild-type donor (IV:10) as compared by Student t test.

C20orf7 is less essential for proper assembly or the defect may be partially rescued in our patients by its absence in isoform 2. As no difference was observed in the relative expression of the two isoforms in different tissues, including brain, liver, heart, skeletal muscle, kidney, adipose tissue, thyroid, blood and endothelium (data not shown), it is unlikely that tissue specific expression of the two isoforms can explain the variation in clinical expression. It remains unclear whether the second isoform leads to a protein with the same function as isoform 1. Further studies are required to resolve the exact pathological or rescue process occurring.

Web resources

The URLs for data presented herein are as follows:

- ▶ Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/>
- ▶ MitoP2 database, <http://www.mitop.de:8080/mitop2/>
- ▶ NCBI SNP database, <http://www.ncbi.nlm.nih.gov/sites/entrez?db=snp>
- ▶ Genome Variation Server, <http://gvs.gs.washington.edu/GVS/>
- ▶ Nucleotide Variation and Mutation Database MutationDiscovery.com, http://www.mutationdiscovery.com/md/MD.com/home_page.jsp
- ▶ ClustalW, <http://www.ebi.ac.uk/Tools/clustalw2/index.html>
- ▶ SOPM, http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopm.html

Patients were not only homozygous for the mutation in *C20orf7*, but also for a common variant in the *CRLS1* gene. As 1% of the Dutch and 4% of the Moroccan population is predicted to be homozygous for this mutation, this would imply a disease frequency of 1–4 in 100, if the *CRLS1* variant would be the genetic cause. This is not the case. However, the reduction of CLS activity, which is unlikely to cause clinical symptoms by itself, may contribute to the disease phenotype in case the OXPHOS is already affected by another defect, like in this family. So, *C20orf7* mutations are a new cause of Leigh syndrome due to an early complex I assembly defect. Disease manifestations are mutation specific and may be modified by additional mutations in other genes, like *CRLS1*.

Acknowledgements This work was supported by a EU grant to the MitoCircle project (Sixth Framework Program, contr. no. 005260). We thank Prof Dr B Oostra for supplying a Moroccan control panel.

Competing interests None.

Patient consent Obtained.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. Dahl HH. Getting to the nucleus of mitochondrial disorders: identification of respiratory chain-enzyme genes causing Leigh syndrome. *Am J Hum Genet* 1998;**63**:1594–7.
2. Rotig A, Munnich A. Genetic features of mitochondrial respiratory chain disorders. *J Am Soc Nephrol* 2003;**14**:2995–3007.
3. Moslemi AR, Darin N, Tulinius M, Wiklund LM, Holme E, Oldfors A. Progressive encephalopathy and complex I deficiency associated with mutations in *MTND1*. *Neuropediatrics* 2008;**39**:24–8.
4. Hinttala R, Smeets R, Moilanen JS, Ugalde C, Uusimaa J, Smeitink JA, Majamaa K. Analysis of mitochondrial DNA sequences in patients with isolated or combined oxidative phosphorylation system deficiency. *J Med Genet* 2006;**43**:881–6.
5. McFarland R, Kirby DM, Fowler KJ, Ohtake A, Ryan MT, Amor DJ, Fletcher JM, Dixon JW, Collins FA, Turnbull DM, Taylor RW, Thorburn DR. De novo mutations in the mitochondrial *ND3* gene as a cause of infantile mitochondrial encephalopathy and complex I deficiency. *Ann Neurol* 2004;**55**:58–64.
6. Taylor RW, Morris AA, Hutchinson M, Turnbull DM. Leigh disease associated with a novel mitochondrial DNA *ND5* mutation. *Eur J Hum Genet* 2002;**10**:141–4.
7. Ugalde C, Triepels RH, Coenen MJ, van den Heuvel LP, Smeets R, Uusimaa J, Briones P, Campistol J, Majamaa K, Smeitink JA, Nijtmans LG. Impaired enzymatic assembly in a Leigh syndrome patient with a novel missense mutation in the *ND6* gene. *Ann Neurol* 2003;**54**:665–9.
8. Schuelke M, Smeitink J, Mariman E, Loeffen J, Plecko B, Trijbels F, Stöckler-Ipsiroglu S, van den Heuvel L. Mutant *NDUF1* subunit of mitochondrial complex I causes leukodystrophy and myoclonic epilepsy. *Nat Genet* 1999;**21**:260–1.
9. Benit P, Chretien D, Kadhom N, de Lonlay-Debeney P, Cormier-Daire V, Cabral A, Peudenier S, Rustin P, Munnich A, Rötig A. Large-scale deletion and point mutations of the nuclear *NDUFV1* and *NDUFS1* genes in mitochondrial complex I deficiency. *Am J Hum Genet* 2001;**68**:1344–52.
10. Benit P, Slama A, Cartault F, Giurgea I, Chretien D, Lebon S, Marsac C, Munnich A, Rötig A, Rustin P. Mutant *NDUFS3* subunit of mitochondrial complex I causes Leigh syndrome. *J Med Genet* 2004;**41**:14–17.
11. Budde SM, van den Heuvel LP, Janssen AJ, Smeets RJ, Buskens CA, DeMeirleir L, Van Coster R, Baethmann M, Voit T, Trijbels JM, Smeitink JA. Combined enzymatic complex I and III deficiency associated with mutations in the nuclear encoded *NDUFS4* gene. *Biochem Biophys Res Commun* 2000;**275**:63–8.
12. Lebon S, Rodriguez D, Bridoux D, Zerrad A, Rötig A, Munnich A, Legrand A, Slama A. A novel mutation in the human complex I *NDUFS7* subunit associated with Leigh syndrome. *Mol Genet Metab* 2007;**90**:379–82.
13. Loeffen J, Smeitink J, Triepels R, Smeets R, Schuelke M, Sengers R, Trijbels F, Hamel B, Mullaart R, van den Heuvel L. The first nuclear-encoded complex I mutation in a patient with Leigh syndrome. *Am J Hum Genet* 1998;**63**:1598–608.
14. Berger I, Hershkovitz E, Shaag A, Edvardson S, Saada A, Elpeleg O. Mitochondrial complex I deficiency caused by a deleterious *NDUFA11* mutation. *Ann Neurol* 2008;**63**:405–8.
15. Hoefs SJ, Dieteren CE, Distelmaier F, Janssen RJ, Epplen A, Swarts HG, Forkink M, Rodenburg RJ, Nijtmans LG, Willems PH, Smeitink JA, van den Heuvel LP. *NDUFA2* complex I mutation leads to Leigh disease. *Am J Hum Genet* 2008;**82**:1306–15.
16. Bourgeron T, Rustin P, Chretien D, Birch-Machin M, Bourgeois M, Viegas-Péguignot E, Munnich A, Rötig A. Mutation of a nuclear succinate dehydrogenase gene results in mitochondrial respiratory chain deficiency. *Nat Genet* 1995;**11**:144–9.

17. **de Lonlay P**, Valnot I, Barrientos A, Gorbatyuk M, Tzagoloff A, Taanman JW, Benayoun E, Chrétien D, Kadhon N, Lombès A, de Baulny HO, Niaudet P, Munnich A, Rustin P, Rötig A. A mutant mitochondrial respiratory chain assembly protein causes complex III deficiency in patients with tubulopathy, encephalopathy and liver failure. *Nat Genet* 2001;**29**:57–60.
18. **Keightley JA**, Hoffbuhr KC, Burton MD, Salas VM, Johnston WS, Penn AM, Buist NR, Kennaway NG. A microdeletion in cytochrome c oxidase (COX) subunit III associated with COX deficiency and recurrent myoglobinuria. *Nat Genet* 1996;**12**:410–16.
19. **Valnot I**, von Kleist-Retzow JC, Barrientos A, Gorbatyuk M, Taanman JW, Mehaye B, Rustin P, Tzagoloff A, Munnich A, Rötig A. A mutation in the human heme A: farnesyltransferase gene (COX10) causes cytochrome c oxidase deficiency. *Hum Mol Genet* 2000;**9**:1245–9.
20. **Oquendo CE**, Antonicka H, Shoubridge EA, Reardon W, Brown GK. Functional and genetic studies demonstrate that mutation in the COX15 gene can cause Leigh syndrome. *J Med Genet* 2004;**41**:540–4.
21. **Papadopoulou LC**, Sue CM, Davidson MM, Tanji K, Nishino I, Sadlock JE, Krishna S, Walker W, Selby J, Glerum DM, Coster RV, Lyon G, Scalais E, Lebel R, Kaplan P, Shanske S, De Vivo DC, Bonilla E, Hirano M, DiMauro S, Schon EA. Fatal infantile cardioencephalomyopathy with COX deficiency and mutations in SCO2, a COX assembly gene. *Nat Genet* 1999;**23**:333–7.
22. **Tiranti V**, Hoertnagel K, Carozzo R, Galimberti C, Munaro M, Granatiero M, Zelante L, Gasparini P, Marzella R, Rocchi M, Bayona-Bafaluy MP, Enriquez JA, Uziel G, Bertini E, Dionisi-Vici C, Franco B, Meitinger T, Zeviani M. Mutations of SURF-1 in Leigh disease associated with cytochrome c oxidase deficiency. *Am J Hum Genet* 1998;**63**:1609–21.
23. **de Coo IF**, Smeets HJ, Gabreels FJ, Arts N, van Oost BA. Isolated case of mental retardation and ataxia due to a de novo mitochondrial T8993G mutation. *Am J Hum Genet* 1996;**58**:636–8.
24. **McFarland R**, Clark KM, Morris AA, Taylor RW, Macphail S, Lightowlers RN, Turnbull DM. Multiple neonatal deaths due to a homoplasmic mitochondrial DNA mutation. *Nat Genet* 2002;**30**:145–6.
25. **Shoffner JM**, Lott MT, Lezza AM, Seibel P, Ballinger SW, Wallace DC. Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a mitochondrial DNA tRNA(Lys) mutation. *Cell* 1990;**61**:931–7.
26. **Santorelli FM**, Tanji K, Sano M, Shanske S, El-Shahawi M, Kranz-Eble P, DiMauro S, De Vivo DC. Maternally inherited encephalopathy associated with a single-base insertion in the mitochondrial tRNA^{Trp} gene. *Ann Neurol* 1997;**42**:256–60.
27. **Kobayashi Y**, Momoi MY, Tominaga K, Momoi T, Nihei K, Yanagisawa M, Kagawa Y, Ohta S. A point mutation in the mitochondrial tRNA(Leu)(UUR) gene in MELAS (mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes). *Biochem Biophys Res Commun* 1990;**173**:816–22.
28. **Grafakou O**, Oexle K, van den Heuvel L, Smeets R, Trijbels F, Goebel HH, Bosshard N, Superti-Furga A, Steinmann B, Smeitink J. Leigh syndrome due to compound heterozygosity of dihydrolipoamide dehydrogenase gene mutations. Description of the first E3 splice site mutation. *Eur J Pediatr* 2003;**162**:714–18.
29. **Matthews PM**, Marchington DR, Squier M, Land J, Brown RM, Brown GK. Molecular genetic characterization of an X-linked form of Leigh's syndrome. *Ann Neurol* 1993;**33**:652–5.
30. **Mootha VK**, Lepage P, Miller K, Bunkenborg J, Reich M, Hjerrild M, Delmonte T, Villeneuve A, Sladek R, Xu F, Mitchell GA, Morin C, Mann M, Hudson TJ, Robinson B, Rioux JD, Lander ES. Identification of a gene causing human cytochrome c oxidase deficiency by integrative genomics. *Proc Natl Acad Sci U S A* 2003;**100**:605–10.
31. **Sugiana C**, Pagliarini DJ, McKenzie M, Kirby DM, Salemi R, Abu-Amero KK, Dahl HH, Hutchison WM, Vascotto KA, Smith SM, Newbold RF, Christodoulou J, Calvo S, Mootha VK, Ryan MT, Thorburn DR. Mutation of C20orf7 disrupts complex I assembly and causes lethal neonatal mitochondrial disease. *Am J Hum Genet* 2008;**83**:468–78.
32. **de Wit LE**, Scholte HR, Sluiter W. Correct assay of complex I activity in human skin fibroblasts by timely addition of rotenone. *Clin Chem* 2008;**54**:1921–2; author reply 1922–4.
33. **de Wit LE**, Spruijt L, Schoonderwoerd GC, de Coo IF, Smeets HJ, Scholte HR, Sluiter W. A simplified and reliable assay for complex I in human blood lymphocytes. *J Immunol Methods* 2007;**326**:76–82.
34. **de Wit LE**, Sluiter W. Chapter 9 Reliable assay for measuring complex I activity in human blood lymphocytes and skin fibroblasts. *Methods Enzymol* 2009;**456**:169–81.
35. **Vydt TC**, de Coo RF, Soliman OI, Ten Cate FJ, van Geuns RJ, Vletter WB, Schoonderwoerd K, van den Bosch BJ, Smeets HJ, Geleijnse ML. Cardiac involvement in adults with m.3243A > G MELAS gene mutation. *Am J Cardiol* 2007;**99**:264–9.
36. **Nowicki M**, Muller F, Frentzen M. Cardiolipin synthase of Arabidopsis thaliana. *FEBS Lett* 2005;**579**:2161–5.
37. **Abecasis GR**, Cherny SS, Cookson WO, Cardon LR. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002;**30**:97–101.
38. **van Den Bosch BJ**, de Coo RF, Scholte HR, Nijland JG, van Den Bogaard R, de Visser M, de Die-Smulders CE, Smeets HJ. Mutation analysis of the entire mitochondrial genome using denaturing high performance liquid chromatography. *Nucleic Acids Res* 2000;**28**:E89.
39. **van Raam BJ**, Sluiter W, de Wit E, Roos D, Verhoeven AJ, Kuijpers TW. Mitochondrial membrane potential in human neutrophils is maintained by complex III activity in the absence of supercomplex organisation. *PLoS ONE* 2008;**3**:e2013.
40. **Calvo S**, Jain M, Xie X, Sheth SA, Chang B, Goldberger OA, Spinazzola A, Zeviani M, Carr SA, Mootha VK. Systematic identification of human mitochondrial disease genes through integrative genomics. *Nat Genet* 2006;**38**:576–82.
41. **Combat C**, Blanchet C, Geourjon C, et al. NPS@: network protein sequence analysis. *Trends Biochem Sci* 2000;**25**:147–50.
42. **Lazarou M**, Thorburn DR, Ryan MT, McKenzie M. Assembly of mitochondrial complex I and defects in disease. *Biochim Biophys Acta* 2009;**1793**:78–88.

Table 1. PCR primers *C20orf7* and *CRLS1*.

PCR primers 5' → 3'			
Gene-Exon	Forward	Reverse	Product size (bp)
C20orf7-Ex1	GCACGTTCCCTCGAGAAGAT	TAGCGGAGAGGAAAACAAGC	379
C20orf7-Ex2	TGTTTCCTGGTGGTTTGAGA	AGTGGATTCAGGCTTGGAGA	442
C20orf7-Ex3	AAGCTGGCTCATCACAAGGT	TGTTGCAAGCTAAACTGATTAGGA	371
C20orf7-Ex4	TTGAATACGCATACTCATGTTCTC	TGATATTTCTGTATGAGTTCCTTTT	400
C20orf7-Ex5	CCTTCTCTGTGCCACCTTTG	TCCTGACCTCAAGAGATTCACTC	391
C20orf7-Ex6	ACGATCAGTGGATTTGGATATG	TGCACAGTTCCTAAGCTTTTCA	423
C20orf7-Ex7	TGATTTTCTGACAGTTGGTTGC	TGCCATGAAACCATGAACAC	471
C20orf7-Ex8	TGATTTCTCTCACAATGCAGGA	CTCTGTGACAGCTGCCTGGT	484
C20orf7-Ex9	TCGCTTTAGAGTCACTAAGAAACC	TGAGGGAGCCTTAACAGGAA	372
C20orf7-Ex10-11	TCCTTG TACCCATTTCTTTTTGA	GCTTAGAGTGCTTCTTGCTTCA	533
CRLS1-Ex1	AGCCCAGTGTCTGAGTGGTT	CCACCCAGAGCTTGTTACCT	493
CRLS1-Ex2	TGTAGCTTCTGACTGAAGTTGATG	CAAGGAAAAGCAGGCTAAGG	343
CRLS1-Ex3	TGCTTTGTTCTGTTATGACATTT	TTGTCTCTCAAGAGAAATCCTAAGC	264
CRLS1-Ex4	AGCAGTTTGTGGTAAGGCTGA	TCTTTTTGCTTTTTCTTAAATCA	250
CRLS1-Ex5	TCGAATTTTACAGCAATTGAAAA	GTAAAACTGCCTGCCCAAC	400
CRLS1-Ex6	GGATCCAGTTGGCACTATTCA	TCAGTGCTCAACAAATGCTGT	240
CRLS1-Ex7	TTCATTATGATTCCCGTTGC	GCCCATGCTCTGGACTATGT	479

Table S1 Detailed SNP information of the homozygous region

Affymetrix SNP ID	Marker	Chromosome	Physical Position	III:1	III:2	IV:1	IV:5	IV:6	IV:10	IV:7	IV:11
SNP_A-1512974	rs2013961	20	2265490	BB	AA	AB	AB	AB	AB	AB	AB
SNP_A-1511055	rs1109010	20	2405887	BB	BB	BB	BB	BB	BB	BB	BB
SNP_A-1512375	rs3848810	20	3522148	AA	AA	AA	AA	AA	AA	AA	AA
SNP_A-1509541	rs910652	20	3675970	AA	AB	AB	AA	AA	AB	AA	AA
SNP_A-1509668	rs674110	20	3676026	BB	BB	BB	BB	BB	BB	BB	BB
SNP_A-1517452	rs910952	20	4141694	BB	BB	BB	BB	BB	BB	BB	BB
SNP_A-1510407	rs2422925	20	4251877	AA	AB	AB	AA	AA	AB	AA	AA
SNP_A-1513801	rs3904872	20	4260153	NoCall	NoCall	NoCall	BB	BB	NoCall	NoCall	NoCall
SNP_A-1512173	rs3904864	20	4327086	AA	AB	AB	AA	AA	AB	AA	AA
SNP_A-1511202	rs1411296	20	6437010	AA	AB	AB	AA	AA	AB	AA	AA
SNP_A-1508028	rs723131	20	6529919	NoCall	AB	NoCall	AB	AB	AA	BB	BB
SNP_A-1512742	rs2224191	20	6755276	AA	AA	AA	AA	AA	AA	AA	AA
SNP_A-1508861	rs723511	20	6786358	AB	AA	AA	AB	AB	AB	AA	AA
SNP_A-1515430	rs717087	20	6862914	BB	BB	BB	BB	BB	BB	BB	BB
SNP_A-1519528	rs952793	20	6918831	AB	AB	AB	AB	AB	BB	AA	AA
SNP_A-1517385	rs2149642	20	6962445	AB	AB	AB	AB	AB	BB	AA	AA
SNP_A-1513440	rs717074	20	7120538	AA	AA	AA	AA	AA	AA	AA	AA
SNP_A-1513450	rs720010	20	7174248	AA	AA	AA	AA	AA	AA	AA	AA
SNP_A-1518624	rs2009752	20	7188177	AA	AA	AA	NoCall	AA	AA	AA	AA
SNP_A-1513356	rs724084	20	7223683	BB	BB	BB	BB	BB	BB	BB	BB
SNP_A-1513218	rs2224053	20	7254965	AB	AB	AB	AB	AB	BB	AA	AA
SNP_A-1510348	rs2223276	20	7830106	BB	BB	BB	BB	BB	BB	BB	BB
SNP_A-1515069	rs2250711	20	8203869	BB	BB	NoCall	BB	BB	BB	BB	BB
SNP_A-1509024	rs813467	20	8361676	AB	AB	AB	AB	AB	BB	AA	AA
SNP_A-1507452	rs771943	20	8362876	AB	AB	AB	AB	AB	AA	BB	BB
SNP_A-1518940	rs1473685	20	8411131	AA	AB	AB	AA	AA	AB	AA	AA

SNP_A-1514966	rs2223539	20	8489270	AB	AA	AA	AB	AB	AB	AA	AA
SNP_A-1519337	rs2206489	20	8714796	BB	AB	AB	BB	BB	AB	BB	BB
SNP_A-1512276	rs2327088	20	8714882	AB	BB	BB	AB	AB	AB	BB	BB
SNP_A-1515630	rs2143267	20	8757225	BB	BB	BB	BB	BB	BB	BB	BB
SNP_A-1517103	rs3848831	20	8799321	AA	AB	AB	AA	AA	AB	AA	AA
SNP_A-1513028	rs756374	20	8892142	AB	AB	AB	AB	AB	AA	BB	BB
SNP_A-1514406	rs724089	20	9001556	AB	AA	AA	AB	AB	AB	AA	AA
SNP_A-1516905	rs953021	20	9898715	BB	AB	AB	BB	BB	AB	BB	BB
SNP_A-1514455	rs725565	20	9953010	AB	BB	BB	AB	AB	AB	BB	BB
SNP_A-1509328	rs2327282	20	10375960	AB	AB	AB	AB	AB	AA	BB	BB
SNP_A-1510576	rs973542	20	10640559	AB	BB	BB	AB	AB	AB	BB	BB
SNP_A-1519560	rs2327302	20	10663404	BB	AB	AB	BB	BB	AB	BB	BB
SNP_A-1509521	rs763383	20	10845878	AB	AB	AB	AB	AB	BB	AA	AA
SNP_A-1514736	rs1028846	20	10960899	AB	BB	BB	AB	AB	AB	BB	BB
SNP_A-1512494	rs726867	20	11175697	AB	AB	AB	AB	AB	BB	AA	AA
SNP_A-1508024	rs3887413	20	11754851	AB	BB	BB	AB	AB	AB	BB	BB
SNP_A-1508766	rs803876	20	11798158	AA	AA	AA	AA	AA	AA	AA	AA
SNP_A-1511241	rs2206798	20	11845445	BB	AB	AB	BB	BB	AB	BB	BB
SNP_A-1517499	rs720489	20	11897114	AA	AA	AA	AA	AA	AA	AA	AA
SNP_A-1510612	rs729552	20	12131209	AB	BB	BB	AB	AB	AB	BB	BB
SNP_A-1509951	rs2208970	20	12325493	AA	AA	AA	AA	AA	AA	AA	AA
SNP_A-1513642	rs912086	20	12339974	BB	BB	BB	BB	BB	BB	NoCall	NoCall
SNP_A-1514630	rs1074071	20	12633208	BB	BB	BB	BB	BB	BB	BB	BB
SNP_A-1515199	rs1099620	20	12738724	AB	AB	AB	AB	AB	BB	AA	AA
SNP_A-1509681	rs3908399	20	12849275	AB	AA	AA	AB	AB	AB	AA	AA
SNP_A-1511747	rs2327450	20	12925506	BB	AB	AB	BB	BB	AB	BB	BB
SNP_A-1508911	rs721243	20	13168601	AB	BB	BB	AB	AB	AB	BB	BB
SNP_A-1514723	rs719628	20	13210487	AB	AB	AB	AB	AB	AA	BB	BB

SNP_A-1513229	rs2327790	20	13461712	AB	BB	BB	AB	AB	AB	BB	BB
C20orf7											
SNP_A-1519307	rs1358766	20	13858933	AB	AB	AB	AB	AB	BB	AA	AA
SNP_A-1513470	rs724820	20	15015407	AB	AB	AA	BB	BB	AB	AB	AB

Overview of genotypes called by GDAS for the SNPs located in the homozygous region. For each marker the Affy-ID and physical position is indicated. Sample IDs refer to family IDs in figure1.

Table 1. PCR primers *C20orf7* and *CRLS1*.

PCR primers 5' → 3'			
Gene-Exon	Forward	Reverse	Product size (bp)
C20orf7-Ex1	GCACGTTCCCTCGAGAAGAT	TAGCGGAGAGGAAAACAAGC	379
C20orf7-Ex2	TGTTTCCTGGTGGTTTGAGA	AGTGGATTCAGGCTTGGAGA	442
C20orf7-Ex3	AAGCTGGCTCATCACAAGGT	TGTTGCAAGCTAAACTGATTAGGA	371
C20orf7-Ex4	TTGAATACGCATACTCATGTTCTC	TGATATTTCTGTATGAGTTCCTTTT	400
C20orf7-Ex5	CCTTCTCTGTGCCACCTTTG	TCCTGACCTCAAGAGATTCACTC	391
C20orf7-Ex6	ACGATCAGTGGATTTGGATATG	TGCACAGTTCCTAAGCTTTTCA	423
C20orf7-Ex7	TGATTTTCTGACAGTTGGTTGC	TGCCATGAAACCATGAACAC	471
C20orf7-Ex8	TGATTTCTCTCACAATGCAGGA	CTCTGTGACAGCTGCCTGGT	484
C20orf7-Ex9	TCGCTTTAGAGTCACTAAGAAACC	TGAGGGAGCCTTAACAGGAA	372
C20orf7-Ex10-11	TCCTTG TACCCATTTCTTTTTGA	GCTTAGAGTGCTTCTTGCTTCA	533
CRLS1-Ex1	AGCCCAGTGTCTGAGTGGTT	CCACCCAGAGCTTGTTACCT	493
CRLS1-Ex2	TGTAGCTTCTGACTGAAGTTGATG	CAAGGAAAAGCAGGCTAAGG	343
CRLS1-Ex3	TGCTTTGTTCTGTTATGACATTT	TTGTCTCTCAAGAGAAATCCTAAGC	264
CRLS1-Ex4	AGCAGTTTGTGGTAAGGCTGA	TCTTTTTGCTTTTTCTTAAATCA	250
CRLS1-Ex5	TCGAATTTTACAGCAATTGAAAA	GTAAAACTGCCTGCCCAAC	400
CRLS1-Ex6	GGATCCAGTTGGCACTATTCA	TCAGTGCTCAACAAATGCTGT	240
CRLS1-Ex7	TTCATTATGATTCCCGTTGC	GCCCATGCTCTGGACTATGT	479

Table S1 Detailed SNP information of the homozygous region

Affymetrix SNP ID	Marker	Chromosome	Physical Position	III:1	III:2	IV:1	IV:5	IV:6	IV:10	IV:7	IV:11
SNP_A-1512974	rs2013961	20	2265490	BB	AA	AB	AB	AB	AB	AB	AB
SNP_A-1511055	rs1109010	20	2405887	BB	BB	BB	BB	BB	BB	BB	BB
SNP_A-1512375	rs3848810	20	3522148	AA	AA	AA	AA	AA	AA	AA	AA
SNP_A-1509541	rs910652	20	3675970	AA	AB	AB	AA	AA	AB	AA	AA
SNP_A-1509668	rs674110	20	3676026	BB	BB	BB	BB	BB	BB	BB	BB
SNP_A-1517452	rs910952	20	4141694	BB	BB	BB	BB	BB	BB	BB	BB
SNP_A-1510407	rs2422925	20	4251877	AA	AB	AB	AA	AA	AB	AA	AA
SNP_A-1513801	rs3904872	20	4260153	NoCall	NoCall	NoCall	BB	BB	NoCall	NoCall	NoCall
SNP_A-1512173	rs3904864	20	4327086	AA	AB	AB	AA	AA	AB	AA	AA
SNP_A-1511202	rs1411296	20	6437010	AA	AB	AB	AA	AA	AB	AA	AA
SNP_A-1508028	rs723131	20	6529919	NoCall	AB	NoCall	AB	AB	AA	BB	BB
SNP_A-1512742	rs2224191	20	6755276	AA	AA	AA	AA	AA	AA	AA	AA
SNP_A-1508861	rs723511	20	6786358	AB	AA	AA	AB	AB	AB	AA	AA
SNP_A-1515430	rs717087	20	6862914	BB	BB	BB	BB	BB	BB	BB	BB
SNP_A-1519528	rs952793	20	6918831	AB	AB	AB	AB	AB	BB	AA	AA
SNP_A-1517385	rs2149642	20	6962445	AB	AB	AB	AB	AB	BB	AA	AA
SNP_A-1513440	rs717074	20	7120538	AA	AA	AA	AA	AA	AA	AA	AA
SNP_A-1513450	rs720010	20	7174248	AA	AA	AA	AA	AA	AA	AA	AA
SNP_A-1518624	rs2009752	20	7188177	AA	AA	AA	NoCall	AA	AA	AA	AA
SNP_A-1513356	rs724084	20	7223683	BB	BB	BB	BB	BB	BB	BB	BB
SNP_A-1513218	rs2224053	20	7254965	AB	AB	AB	AB	AB	BB	AA	AA
SNP_A-1510348	rs2223276	20	7830106	BB	BB	BB	BB	BB	BB	BB	BB
SNP_A-1515069	rs2250711	20	8203869	BB	BB	NoCall	BB	BB	BB	BB	BB
SNP_A-1509024	rs813467	20	8361676	AB	AB	AB	AB	AB	BB	AA	AA
SNP_A-1507452	rs771943	20	8362876	AB	AB	AB	AB	AB	AA	BB	BB
SNP_A-1518940	rs1473685	20	8411131	AA	AB	AB	AA	AA	AB	AA	AA

SNP_A-1514966	rs2223539	20	8489270	AB	AA	AA	AB	AB	AB	AA	AA
SNP_A-1519337	rs2206489	20	8714796	BB	AB	AB	BB	BB	AB	BB	BB
SNP_A-1512276	rs2327088	20	8714882	AB	BB	BB	AB	AB	AB	BB	BB
SNP_A-1515630	rs2143267	20	8757225	BB	BB	BB	BB	BB	BB	BB	BB
SNP_A-1517103	rs3848831	20	8799321	AA	AB	AB	AA	AA	AB	AA	AA
SNP_A-1513028	rs756374	20	8892142	AB	AB	AB	AB	AB	AA	BB	BB
SNP_A-1514406	rs724089	20	9001556	AB	AA	AA	AB	AB	AB	AA	AA
SNP_A-1516905	rs953021	20	9898715	BB	AB	AB	BB	BB	AB	BB	BB
SNP_A-1514455	rs725565	20	9953010	AB	BB	BB	AB	AB	AB	BB	BB
SNP_A-1509328	rs2327282	20	10375960	AB	AB	AB	AB	AB	AA	BB	BB
SNP_A-1510576	rs973542	20	10640559	AB	BB	BB	AB	AB	AB	BB	BB
SNP_A-1519560	rs2327302	20	10663404	BB	AB	AB	BB	BB	AB	BB	BB
SNP_A-1509521	rs763383	20	10845878	AB	AB	AB	AB	AB	BB	AA	AA
SNP_A-1514736	rs1028846	20	10960899	AB	BB	BB	AB	AB	AB	BB	BB
SNP_A-1512494	rs726867	20	11175697	AB	AB	AB	AB	AB	BB	AA	AA
SNP_A-1508024	rs3887413	20	11754851	AB	BB	BB	AB	AB	AB	BB	BB
SNP_A-1508766	rs803876	20	11798158	AA	AA	AA	AA	AA	AA	AA	AA
SNP_A-1511241	rs2206798	20	11845445	BB	AB	AB	BB	BB	AB	BB	BB
SNP_A-1517499	rs720489	20	11897114	AA	AA	AA	AA	AA	AA	AA	AA
SNP_A-1510612	rs729552	20	12131209	AB	BB	BB	AB	AB	AB	BB	BB
SNP_A-1509951	rs2208970	20	12325493	AA	AA	AA	AA	AA	AA	AA	AA
SNP_A-1513642	rs912086	20	12339974	BB	BB	BB	BB	BB	BB	NoCall	NoCall
SNP_A-1514630	rs1074071	20	12633208	BB	BB	BB	BB	BB	BB	BB	BB
SNP_A-1515199	rs1099620	20	12738724	AB	AB	AB	AB	AB	BB	AA	AA
SNP_A-1509681	rs3908399	20	12849275	AB	AA	AA	AB	AB	AB	AA	AA
SNP_A-1511747	rs2327450	20	12925506	BB	AB	AB	BB	BB	AB	BB	BB
SNP_A-1508911	rs721243	20	13168601	AB	BB	BB	AB	AB	AB	BB	BB
SNP_A-1514723	rs719628	20	13210487	AB	AB	AB	AB	AB	AA	BB	BB

SNP_A-1513229	rs2327790	20	13461712	AB	BB	BB	AB	AB	AB	BB	BB
C20orf7											
SNP_A-1519307	rs1358766	20	13858933	AB	AB	AB	AB	AB	BB	AA	AA
SNP_A-1513470	rs724820	20	15015407	AB	AB	AA	BB	BB	AB	AB	AB

Overview of genotypes called by GDAS for the SNPs located in the homozygous region. For each marker the Affy-ID and physical position is indicated. Sample IDs refer to family IDs in figure1.