Bi-allelic Variants in *TKFC* Encoding Triokinase/FMN Cyclase Are Associated with Cataracts and Multisystem Disease

Saskia B. Wortmann,[1,2,3,13] Brigitte Meunier,[4,13] Lamia Mestek-Boukhibar,[5,14] Florence van den Broek,[1,14] Elaina M. Maldonado,[6] Emma Clement,[7] Daniel Weghuber,[1] Johannes Spenger,[1] Zdenek Jaros,[8] Fatma Taha,[6] Wyatt W. Yue,[9] Simon J. Heales,[6,10,11] James E. Davison,[12] Johannes A. Mayr,[1,15] and Shamima Rahman[6,12,15,*]

We report an inborn error of metabolism caused by TKFC deficiency in two unrelated families. Rapid trio genome sequencing in family 1 and exome sequencing in family 2 excluded known genetic etiologies, and further variant analysis identified rare homozygous variants in *TKFC*. *TKFC* encodes a bifunctional enzyme involved in fructose metabolism through its glyceraldehyde kinase activity and in the generation of riboflavin cyclic 4′,5′-phosphate (cyclic FMN) through an FMN lyase domain. The *TKFC* homozygous variants reported here are located within the FMN lyase domain. Functional assays in yeast support the deleterious effect of these variants on protein function. Shared phenotypes between affected individuals with TKFC deficiency include cataracts and developmental delay, associated with cerebellar hypoplasia in one case. Further complications observed in two affected individuals included liver dysfunction and microcytic anemia, while one had fatal cardiomyopathy with lactic acidosis following a febrile illness. We postulate that deficiency of TKFC causes disruption of endogenous fructose metabolism leading to generation of by-products that can cause cataract. In line with this, an affected individual had mildly elevated urinary galactitol, which has been linked to cataract development in the galactosemias. Further, in light of a previously reported role of TKFC in regulating innate antiviral immunity through suppression of MDA5, we speculate that deficiency of TKFC leads to impaired innate immunity in response to viral illness, which may explain the fatal illness observed in the most severely affected individual.

Undiagnosed complex multisystem genetic disorders presenting to the pediatric intensive care unit pose a management conundrum. The need to identify (or exclude) a treatable underlying cause is of paramount importance and is a critical factor dictating management decisions. In recent years, next generation sequencing has increasingly been adopted in critically ill children,[1] both to identify or exclude known treatable disorders as well as to discover novel disease-causing genes. Here we demonstrate the utility of rapid genome sequencing followed by functional studies to identify deficiency of TKFC, a bifunctional enzyme involved in fructose metabolism and the generation of riboflavin cyclic 4′,5′-phosphate (cyclic flavin mononucleotide, cFMN)[2] as a cause of a fatal infantile multisystem disorder.

Affected individual 1-1 (P1-1) was the first child born to consanguineous (first cousin) parents of Gujarati ancestry at term. Concerns about eye movements led to the identification of bilateral congenital nuclear cataracts at 5 months of age. Bilateral lensectomy was performed. Subsequently she had developmental delay, with evidence of motor delay (cruising at 26 months, walking independently at 36 months but with a wide-based gait) and speech delay (babbling at 26 months, first words at 4 years, slowly improving understanding). Initial hearing assessments were normal, although there was possible low-frequency hearing impairment at 4 years of age. She had a simple febrile convulsion at 2 years of age but no other seizures and has continued to make slow developmental progress. General examination revealed microphthalmia but no specific dysmorphic features, and no organomegaly. Magnetic resonance imaging of the brain at 29 months revealed cerebellar hypoplasia, with mild delay of myelin maturation (Figure 1). Cardiac investigations including electro- and echocardiograms were normal and remained normal up to the most recent assessment at 4 years of age. Extensive metabolic investigations revealed no diagnostic abnormalities (Table S1).

[1]University Children's Hospital, Salzburger Landeskliniken (SALK) and Paracelsus Medical University (PMU), 5020 Salzburg, Austria; [2]Institute of Human Genetics, Technical University München, 81675 Munich, Germany; [3]Institute of Human Genetics, Helmholtz Zentrum, 85764 Neuherberg, Germany; [4]Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198 Gif-sur-Yvette, France; [5]GOSgene Centre for Translational Omics and NIHR GOSH Biomedical Research Centre, UCL Great Ormond Street Institute of Child Health, London WC1N 1EH, UK; [6]Mitochondrial Research Group, UCL Great Ormond Street Institute of Child Health, London WC1N 1EH, UK; [7]Department of Clinical Genetics, North East Thames Regional Genetics Service, Great Ormond Street Hospital for Children NHS Trust, London WC1N 3JH, UK; [8]Department of Pediatrics, Landesklinikum Zwettl, 3910 Zwettl, Austria; [9]Structural Genomics Consortium, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7DQ, UK; [10]Neurometabolic Unit, National Hospital for Neurology, Queen Square, London WC1N 3BG, UK; [11]Department of Chemical Pathology, Great Ormond Street Hospital for Children NHS Foundation Trust, London WC1N 3JH, UK; [12]Metabolic Unit, Great Ormond Street Hospital for Children NHS Foundation Trust, London WC1N 3JH, UK
[13]These authors contributed equally to this work
[14]These authors contributed equally to this work
[15]These authors contributed equally to this work
*Correspondence: shamima.rahman@ucl.ac.uk
https://doi.org/10.1016/j.ajhg.2020.01.005.

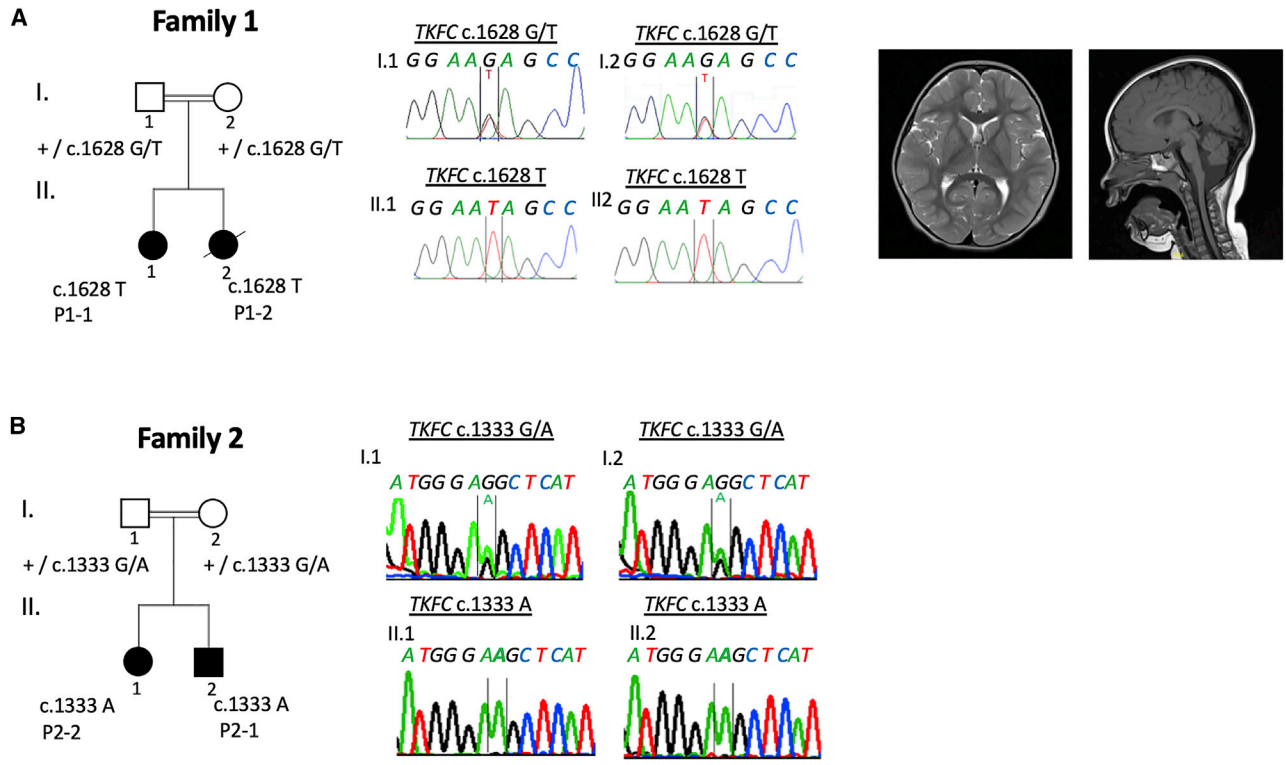


Figure 1. Family Pedigrees, Bi-allelic *TKFC* Variants Observed in Affected Individuals and Brain MRI of P1-1
Representative MRI findings in affected individual P1-1 at 29 months indicating cerebellar hypoplasia. Homozygous *TKFC* variants segregating in each family were confirmed by Sanger Sequencing.

Her younger sister, affected individual 1-2 (P1-2), was diagnosed on day 1 of life with congenital cataracts, which were subsequently excised at 1 month of age. Poor feeding and difficult weight gain were noted from the neonatal period. After her first primary immunizations at 8 weeks, she developed vomiting and loose stool, and was subsequently admitted to hospital, where she was found to be hypotensive and tachypnoeic with hypoglycemia (1.6 mmol/L) and severe lactic acidosis (pH 7.15, lactate 16 mmol/L). Cardiomegaly was evident on a chest radiograph, and echocardiography demonstrated dilated cardiomyopathy with poor systolic function. Liver function was abnormal, with hypoalbuminemia (23 g/L, reference 32–52) and elevated plasma alanine transaminase (900 U/L, reference 10–25). Supportive treatment with intubation, ventilation, inotropic support, and fluid and base therapy led to improvement in the lactic acidosis. Cranial ultrasound demonstrated lenticostriatal vasculopathy, a nonspecific finding which can be seen with viral infections, lactic acidosis, and mitochondrial disease. Magnetic resonance imaging and angiography of the brain was normal at 2 months of age. Attempts at extubation led to deterioration in cardiac function and recurrence of lactic acidosis. Care was redirected and she died aged 11 weeks. Extensive metabolic investigations were unremarkable (Table S1) except for a markedly elevated fibroblast growth factor 21 (FGF21) level of 4,350 pg/mL (reference 44–1,515).

An open muscle biopsy was performed to investigate the possibility of a mitochondrial disorder, since lactic acidaemia, cataracts, cardiomyopathy, and multi-system critical illness are well-recognized features of pediatric mitochondrial diseases. The elevated level of FGF21, which has been postulated as a biomarker of mitochondrial disease,[3] also indicated the possibility of an underlying mitochondrial disorder. Muscle histology did not demonstrate any characteristic features of mitochondrial disease, such as ragged-red or cytochrome oxidase-negative fibers. On electron microscopy, mitochondria appeared slightly enlarged but with a normal morphology. There was a moderate excess of lipid and glycogen content. Blood vessel endothelial and smooth muscle cells contained no atypical inclusions. Respiratory chain enzyme activities in muscle, expressed as a ratio to citrate synthase as a mitochondrial marker enzyme, were essentially within their reference ranges although there was a mild reduction of complex IV activity: complex I 0.251 (reference 0.104–0.268), complex II+III 0.111 (0.04–0.204), and complex IV 0.012 (0.014–0.034).

In view of the critical illness of P1-2 and urgent need to identify any treatable cause of her life-threatening condition, rapid trio genome sequencing of the affected individual and her parents was performed in the RaPS project[1] in tandem with muscle biopsy, array CGH, and mitochondrial DNA (mtDNA) analysis. Array CGH was normal (other than extended regions of homozygosity, as

expected from the pedigree) and mtDNA analysis showed no large-scale rearrangements in blood and normal mtDNA sequence in muscle. Written consent was obtained from the parents of P1-1 and P1-2 to perform trio genome sequencing (of P1-2 and parents) on a research basis under ethics number 08/H0713/82 approved by the NHS Health Research Authority NRES Committee of London Bloomsbury. For research-based whole-genome sequencing (see Supplemental Information for detailed methods), genomic DNA (gDNA) libraries were prepared using Illumina TruSeq DNA PCR-Free Library Prep (Illumina) following the manufacturer's advice starting with 1 μg of sheared gDNA. Parental samples were pooled at equimolar concentrations and sequenced on an Illumina NextSeq 550 in High-Output Mode (29 h). The sample from individual 1-2 was sequenced on an Illumina HighSeq 2500 Dual Flow Cell, Rapid Run Mode (27 h). Mapping and variant calling were performed using a Genalice appliance running Genalice Map 2.5.5 including Mapping, Variant Calling, and the Population Calling module for trio analysis (Genalice Core BV). Genalice default configuration files were used for WGS mapping and trio variant detection. Ingenuity Variant Analysis software (QIAGEN) was used for variant filtering. No potentially pathogenic variants were detected in any known disease genes associated with primary mitochondrial disorders or other inborn errors of metabolism, cardiomyopathy, or intellectual disability. The filtering pipeline prioritized a homozygous variant c.1628G>T (p.Arg543Ile) in *TKFC* (*DAK*, ENSG00000149476, GenBank: NM_015533.3, MIM: 615844) encoding triokinase/FMN cyclase.

Using Genematcher[4] we were able to identify an unrelated family with two affected individuals who also had biallelic variants in *TKFC*, identified by exome sequencing. Affected individual 2-2 (P2-2), a boy, is the second child of consanguineous parents of Turkish ancestry. Pregnancy and birth history, anthropometric data at birth, and the neonatal course were all unremarkable. Poor weight gain and diarrhea were noted from early infancy. Dietary exclusion of cow's milk protein was not associated with any clinical improvement. The affected individual had oral hypersensitivity and did not tolerate introduction of solid foods. At 22 months he was noted to have global developmental delay and ophthalmological evaluation revealed bilateral cataracts. Lens extraction/replacement was performed. Based upon the exome sequencing results, a fructose- and sucrose-free diet was introduced (via gastrostomy in view of oral hypersensitivity) but did not lead to any improvement in weight gain or the diarrhea. Subsequently, the diarrhea worsened and he developed progressive non-cholestatic liver failure with fatty degeneration of the liver. He has been dependent on parenteral feeding since the age of 34 months. His clinical course has been complicated by an episode of pancreatitis. Several attempts to reintroduce significant amounts of enteral feeding failed, and currently he tolerates only 5 × 40 mL of an amino acid-based formula. At 3 years 10 months, weight (10.1 kg) and height (87 cm) are both below the first percentile. He has hepatomegaly but no splenomegaly. He cannot walk independently and has no words. An older sister, affected individual P2-1, has delayed speech development and learning difficulties. She is otherwise well, and ophthalmological examination did not show cataracts. See Table S1 for results of biochemical and hematological investigations. As the affected individual P2-1 is relatively well, she has not had detailed metabolic investigations.

Exome sequencing and variant prioritizing was performed in the index case subject P2-2, his sister P2-1, and parents as reported previously.[5] Studies in family 2 were approved by the Ethics Committee of the Land Salzburg (number 415-E/2552/10-2019). A known homozygous variant c.941C>A (p.Pro314His) in *PAH* (GenBank: NM_000277.1; MIM: 612349) was identified in both children, who were both known to have mild hyperphenylalaninemia identified by newborn screening. This variant has been reported as a cause of benign hyperphenylalaninemia (MIM: 261600)[6] but does not explain the additional phenotypes observed in family 2. We therefore searched for other potentially disease-causing gene variants and a homozygous variant c.1333G>A (p.Gly445Ser) was identified in *TKFC* in P2-1 and P2-2. Biallelic variants in *TKFC* were confirmed by Sanger sequencing and segregated with disease in both families (Figure 1). Genetic evidence for pathogenicity of these variants included the segregation data and the overlapping phenotypes in two unrelated families from distinct geographical regions. *In silico* predictions for both *TKFC* variants are supportive of a deleterious effect (Table 1). The c.1628G>T variant in family 1 is present in the gnomAD database at a very low frequency (2.40E−05) with no homozygous variant reported. The c.1333G>A variant is not present in gnomAD.

TKFC encodes a bifunctional protein that has been annotated as a homodimeric triokinase and FMN cyclase.[2] Triokinase (EC 2.7.1.28) is a component of the fructose metabolism pathway first described by Hers and is responsible for the ATP-dependent phosphorylation of D-glyceraldehyde and exogenous dihydroxyacetone (DHA).[7,8] Both identified variants affect evolutionarily conserved amino acid residues of TKFC (Figure 2A). Site-directed mutagenesis experiments localized amino acid residues 1–339 to the DHA kinase (K) domain and residues 359–575 to the FMN lyase (L) domain, connected by a linker region and assembling into a functional homodimer.[2] Thus, the variants in families 1 and 2 affect the L domain. The variants observed in the affected individuals were modeled on the crystal structure of the *Citrobacter freundii* DHA kinase, which has structural and functional similarities to human TKFC (Figure 2B).[9] The arginine residue at position 543 is located at the surface of the enzyme, and this position is always occupied by a basic residue (almost invariably Arg) among orthologs. Although its precise function is not known, substitution by isoleucine is likely

Table 1. Details of *TKFC* Homozygous Variants Identified in Affected Members of Families 1 and 2

Variant Details	Family 1	Family 2
Position (GRCh37)	11:61113875	11:61112824
Canonical transcript	NM_015533.3	NM_015533.3
cDNA change	c.1628G>T	c.1333G>A
Protein change	p.Arg543Ile	p.Gly445Ser
In Silico Predictions		
SIFT prediction/score	deleterious/0	deleterious/0.01
PolyPhen prediction/score	probably damaging/1	probably damaging/0.99
Mutation taster prediction/score	disease causing/1	disease causing/1
PROVEN prediction/score	damaging/−7.11	damaging/−5.59
Condel prediction/score	deleterious/0.945	deleterious/0.858
CADD PHRED score	32	32
Minor Allele Frequency		
gnomAD	2.40E−05	not in gnomAD
homozygous allele	none	none
SNP/variant accession number	dbSNP: rs547013163	–

to have impact. The glycine at 445 is close to an ATP/flavin adenine dinucleotide (FAD) binding site and substitution to serine may affect ATP/FAD binding.

In order to quantify the enzyme activities, recombinant human TKFC proteins were expressed and purified from *Escherichia coli* (for details see Supplemental Information). TKFC activity was quantified using a coupled spectrophotometric assay and either D-glyceraldehyde or dihydroxyacetone as substrates.[2] As shown in Figure 3A the TKFC activities of p.Gly445Ser and p.Arg543Ser were reduced to 1.92% ± 0.31% and 1.45% ± 0.75% of wild-type, respectively, when using D-glyceraldehyde as substrate. When using dihydroxyacetone as substrate, the relative activities of p.Gly445Ser and p.Arg543Ser were 6.33% ± 1.67% and 6.22% ± 0.04% of wild-type, respectively (Figure 3B). Equal amounts of protein were used for this enzyme activity assay (Figure 3C). In addition, the *TKFC* homozygous variant in affected individuals P1-1 and P1-2 (family 1) results in significantly reduced protein levels in both individuals as indicated in the western blot showing reduced protein content in cultured skin fibroblasts (Figure 3D). Taken together, these results confirm the deleterious effect of these variants on protein function.

TKFC is widely expressed with highest expression noted in the liver and small intestine. This is consistent with a role of TKFC in fructose metabolism, since liver and small intestine are the main fructose-metabolizing tissues.[10] Functional relevance of observed high expression of TKFC in the adrenal glands is less straightforward to explain, but fructose metabolism has been implicated in corticosteroid hormone production.[11]

Since it was not possible to identify a robust phenotype in cells of affected individuals (cultured skin fibroblasts), functional confirmation of pathogenicity of the mutant alleles was performed using yeast models of TKFC deficiency. We utilized the sequence and predicted functional similarity of TKFC to two yeast DHA kinases Dak1 and Dak2[12] (Figure S1) to interrogate the functional significance of the *TKFC* variants identified in the affected individuals presented here. First we tested yeast wild-type and mutants containing a deletion of *DAK1*, *DAK2*, or both *DAK1* and *DAK2*, as well as yeast cells overexpressing *DAK1* or *DAK2*. We observed that overexpression of *DAK1* or *DAK2* in yeast (WT, Δdak1, Δdak2, Δdak1&2) allowed the cells to use DHA as a carbon source, as previously reported.[12] We then expressed wild-type human *TKFC* in yeast and demonstrated that human TKFC allowed the transformed yeast cells to use DHA as carbon source (Figure 4), confirming the functional overlap of human TKFC and yeast Dak1/2. Finally, we tested the effect of the TKFC variants observed in the affected individuals, namely R543I and G445S (R552I and G458S in yeast Dak2), in the yeast model. Yeast cells overexpressing mutant human *TKFC* or mutant yeast *DAK2* failed to grow on DHA as a carbon source (Figure 4). This result thus directly demonstrates the biochemical effect of the variants on TKFC/DAK function, from which we can infer their likely pathogenicity.

The exact function of TKFC remains unclear to a certain extent. The two yeast orthologs Dak1 and Dak2 are expressed in stress conditions, such as heat or osmotic stress, and encode an enzyme involved in detoxification of DHA.[12] A single study investigating the function of this enzyme in humans revealed that it is a bifunctional enzyme with triokinase and FMN cyclase activities,[2] but it is possible that it has other roles. Three inborn errors

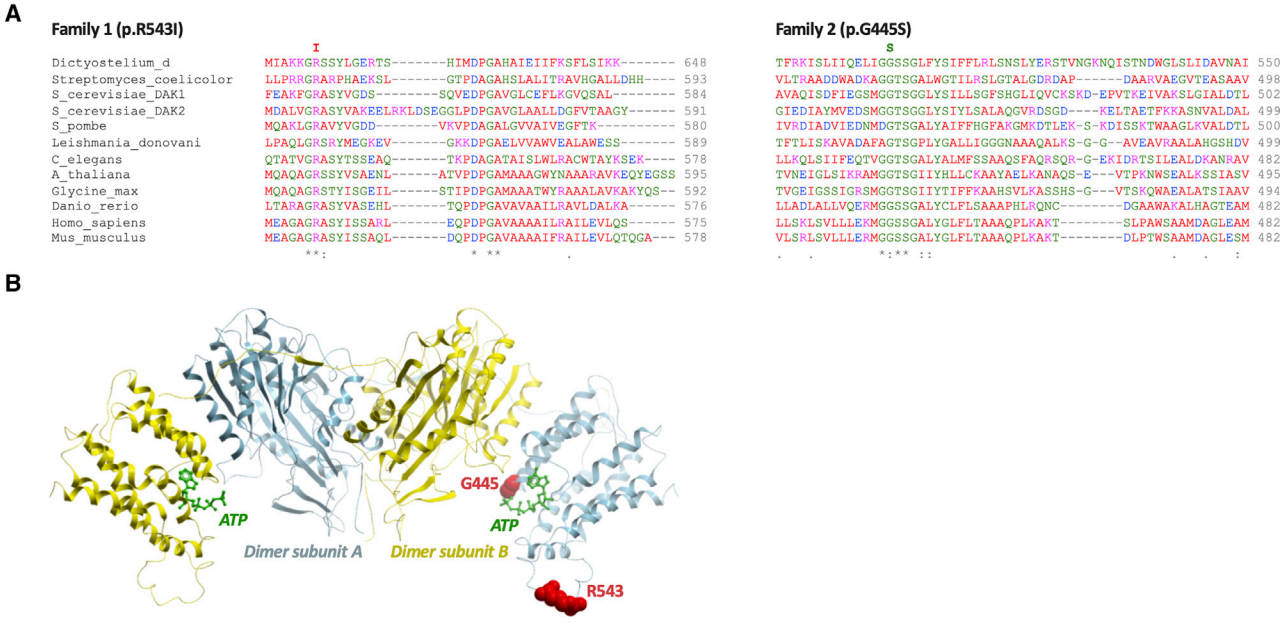


Figure 2. Multiple Sequence Alignment of TKFC Showing Evolutionary Conservation of Affected Amino Acids, and *In Silico* Modeling of TKFC Dimer

(A) Multiple sequence alignment of TKFC among different representative species and location of homozygous variants within TKFC domains. Homozygous protein variants detected in family 1 (p.Arg543Ile) and family 2 (p.Gly445Ser) are indicated and shown to affect highly conserved amino acid residues.

(B) Dimer of TKFC subunits (with K domain shown in yellow and L domain in blue), based on *in silico* modeling of human TKFC structure.[2] The homozygous *TKFC* variants identified in the two families are located within the FMN lyase (L) domain (blue). For simplicity, the two mutation sites (red spheres) are shown only for the dimer subunit A. The ATP ligands from both subunits are represented in green sticks. The p.Gly445Ser variant may moderately affect ATP/FAD binding site due to its close proximity. The p.Arg453Ile variant may have a more severe effect due to predominant preference for only Arg among orthologs.

of fructose metabolism are known: essential fructosuria (fructokinase deficiency [MIM: 229800], a rare non-disease), hereditary fructose intolerance (aldolase B deficiency [MIM: 229600]), and fructose-1,6-bisphosphatase deficiency (MIM: 229700).[13] In hereditary fructose intolerance, fructose ingestion may trigger acute liver failure and proximal renal tubular dysfunction. Fructose-1,6-bisphosphatase deficiency is a disorder of gluconeogenesis and presents with hypoglycaemia and lactic acidosis. The most severely affected individual in our study, P1-2, had not been weaned or exposed to fructose, so it seems unlikely that toxicity from fructose or its metabolites played a significant role in disease pathogenesis. In P2-2 a fructose-free diet did not lead to clinical improvement.

Cataracts were present in three of the four individuals identified to have homozygous *TKFC* variants. Pathogenicity of TKFC deficiency may be by glyceraldehyde formation from impaired fructose catabolism. Glyceraldehyde is a reactive molecule (for instance with hydroxyl and amino groups) and increased production may lead to dysfunction of multiple proteins. Although P1-2 was not exposed to exogenous fructose, she did receive lactose/galactose in milk, so endogenous fructose production via the sorbitol pathway[14] may have contributed to her symptomatology. The observation of mildly increased urinary galactitol (Table S1) in P1-1 is interesting, given the primary role of galactitol in cataract development in the galactosaemias (MIM: 230400, 230200, 230350). The polyol pathway normally metabolizes glucose to sorbitol to fructose, the first step being catalyzed by aldose reductase using NADPH—> NADP+, and the second step by sorbitol dehydrogenase reducing NAD+ to NADH. Galactose is also a substrate for the pathway; aldose reductase can convert galactose to galactitol, but sorbitol dehydrogenase cannot metabolize the galactitol that accumulates, possibly accounting for the cataract formation in affected individuals with TKFC deficiency since aldose reductase is expressed in the lens. The pathway of endogenous fructose production from sorbitol is not well characterized and may be differentially expressed in different tissues and between different individuals, possibly accounting for the phenotypic variability observed in our affected individuals with TKFC deficiency. Another possibility is that the cataracts could be caused by accumulation of DHA due to deficiency in TKFC's DHA kinase activity. Endogenous increase in DHA levels leads to accumulation of advanced glycation end products[15] which are linked to cataract formation.[16,17] Similarly, DHA can itself be a glycation end product or may be metabolized to methylglyoxal, a metabolite previously associated with cataract formation.[18] Additionally, DHA accumulation was shown to induce mitochondrial stress[19] and alter mitochondrial membrane potential leading to apoptotic cell death[20] which might explain the phenotypes mimicking a mitochondrial disorder. Levels of

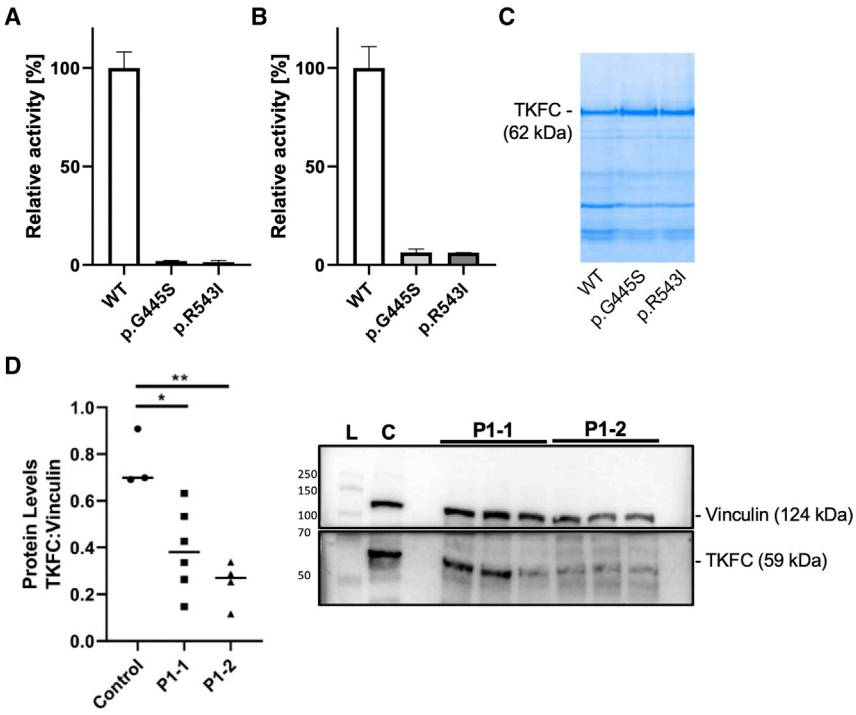


Figure 3. Enzyme Activity of Recombinant Human TKFC Protein Expressed and Purified from *Escherichia coli*, and Western Blot of TKFC in Subject Fibroblasts

(A–C) Enzyme activity was measured using either 10 mmol/L D-glyceraldehyde (A) or 10 mmol/L dihydroxyacetone (B) as substrate. Equal amounts of recombinant protein was used for activity assays as adjusted by polyacrylamide electrophoresis (representative gel in C). Bars show average activity, error bars standard deviation. Recombinant protein was isolated in three replicates in case of wild-type (WT) and the p.Gly445Ser variant and in two replicates for the p.Arg543Ile variant.

(D) TKFC protein studies in family 1 show significant reduction of TKFC protein levels in P1-1 and P1-2 compared to metabolic disease control subjects. Protein levels were calculated as a ratio of TKFC levels to vinculin levels. Data are expressed as median with individual data points. One-way ANOVA with Tukey's multiple comparisons test post hoc was performed to determine significance as indicated by *$p < 0.05$ and **$p < 0.01$. Representative western blot shows lane 1, protein ladder (L); lane 2, metabolic control (C); lanes 4–6, affected individual P1-1; lanes 7–9: affected individual P1-2.

fructose metabolism and endogenous fructose production are cell specific, and metabolism of DHA also appears to be tissue specific.[21]

The critical illness of P1-2 following a febrile illness in early infancy suggests a vital function of TKFC, at least at this developmental stage. The homozygous p.Arg543Ile mutation in this individual affects the cFMN synthase domain. The function of cFMN remains completely unknown, but it has been speculated that it may be a signaling molecule, act as a minor redox flavocoenzyme (although enzymes which might use cFMN as a flavocoenzyme have not been identified)[22] or simply be an intermediate in FAD degradation.[23] TKFC has been shown to be a negative regulator of melanoma differentiation-associated gene-5 (MDA5, encoded by *IFIH1* [MIM: 606951]),[24,25] which is involved in RNA and virus-mediated type I interferon (IFN) production and antiviral responses.[26] Gain-of-function *IFIH1* variants cause heterogeneous phenotypes associated with upregulated type I IFN signaling (MIM: 615846, 182250),[27] while MDA5 deficiency impairs viral double-stranded RNA (dsRNA) sensing and is associated with susceptibility to severe pediatric respiratory syncytial virus and rhinovirus infections in humans.[28,29] Three other genes have been associated with the gene ontology (GO) term "negative regulation of MDA-5 signalling" (GO: 0039534): *C1QBP* (MIM: 601269), *RIOK3* (MIM: 603579), and *DHX58* (MIM: 608588).[30–32] Biallelic *C1QBP* variants in humans cause combined oxidative phosphorylation deficiency 33 (MIM: 617713), a highly variable multisystem mitochondrial disorder with phenotypes ranging from death in infancy to adult-onset progressive external ophthalmoplegia and myopathy. A common finding is cardiomyopathy and increased serum lactate,[33] phenotypes observed in P1-2 in the present study. RIOK3 phosphorylates MDA5, interfering with its assembly and attenuating the innate immune response.[31] DHX58 was shown to function in the RIG-I/MDA5/MAVS protective IFN response in rotavirus-infected intestinal epithelium in mice.[32] The mild loss of complex IV activity observed in the muscle biopsy of P1-2 might reflect secondary damage, for example as a result of oxidative stress following immune response activation. Taking all of these findings into consideration, we postulate that TKFC, via cFMN generation, may modulate innate immune signaling, and that lack of cFMN production in response to a viral insult in a critical developmental window may have led to the fatal illness of P1-2. The TKFC-MDA5 signaling pathway may represent a potential therapeutic target for individuals with TKFC deficiency.

In conclusion, this study demonstrates the power of genome sequencing to exclude a treatable cause of a complex multisystem disorder rapidly in a critical care setting. Extensive metabolic investigation did not reveal any diagnostic clues other than an elevated FGF21 level. This, combined with the clinical features of cataract and cardiomyopathy reminiscent of Sengers syndrome,[5] led to an initial suspicion of mitochondrial disease in P1-2. However, further investigation was not in keeping with a primary mitochondrial disorder. This study demonstrates the utility of genome sequencing and data sharing via Genematcher in the identification of an inborn error of metabolism. Yeast biology is extremely helpful in

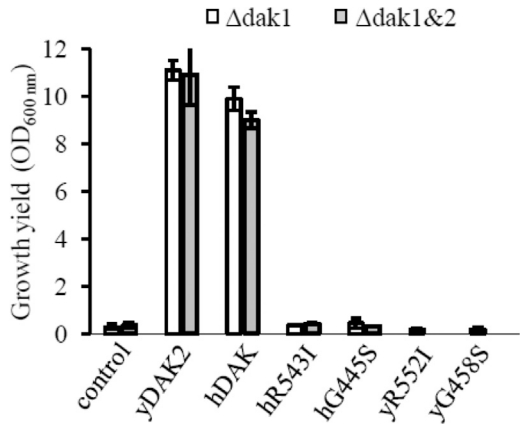


Figure 4. Functional Studies in Yeast
Effect of mutations in wild-type and mutated human and yeast DAK on yeast growth using DHA as sole carbon source. Yeast cells freshly grown on pre-culture plates were inoculated in DHA medium at an initial OD600 nm of 0.2. The cultures were incubated for 4 days at 28°C with vigorous agitation. The OD600 nm were then recorded. The growth experiments were repeated at least twice and the data averaged. Error bars represent standard deviation. Control, cells (Δdak1 or Δdak1&Δdak2) without DAK overexpressing plasmid; yDAK2, cells overexpressing yeast *DAK2*; hDAK, cells overexpressing the WT human DAK/TKFC gene; hR543I and hG445S, cells overexpressing mutated human DAK/TKFC; yR552I and yG458S, cells overexpressing mutated yeast DAK.

dissecting the contribution of variants in housekeeping genes to human disease and here supports the pathogenicity of biallelic *TKFC* variants as a cause of human multisystem disease, variably including cataracts, developmental delay, liver dysfunction, microcytic anemia, cerebellar hypoplasia, and fatal cardiomyopathy with lactic acidosis.

Supplemental Data

Supplemental Data can be found online at https://doi.org/10.1016/j.ajhg.2020.01.005.

Acknowledgments

We thank both families for their participation in this study. Studies in family 1 were funded by National Institute of Health Research (NIHR) Great Ormond Street Hospital (GOSH) Biomedical Research Centre (BRC) NIHR GOSH/UCL BRC: ormbrc-2012-1 and in family 2 by the E-Rare project GENOMIT FWF-I2741B26 for J.A.M. S.R. also receives grant funding from Great Ormond Street Hospital Children's Charity and the Lily Foundation. We are grateful to Drs. Kimberly Gilmour, Elizabeth Ralph, and Mesfer Al Shahrani for performing the FGF21 assays.

Declaration of Interests

The authors declare no competing interests.

Received: November 10, 2019
Accepted: January 7, 2020
Published: January 30, 2020

Web Resources

GenBank, https://www.ncbi.nlm.nih.gov/genbank/
gnomAD Browser, https://gnomad.broadinstitute.org/
GTEx Portal, https://gtexportal.org/home/
OMIM, https://www.omim.org/

References

1. Mestek-Boukhibar, L., Clement, E., Jones, W.D., Drury, S., Ocaka, L., Gagunashvili, A., Le Quesne Stabej, P., Bacchelli, C., Jani, N., Rahman, S., et al. (2018). Rapid Paediatric Sequencing (RaPS): comprehensive real-life workflow for rapid diagnosis of critically ill children. J. Med. Genet. *55*, 721–728.
2. Rodrigues, J.R., Couto, A., Cabezas, A., Pinto, R.M., Ribeiro, J.M., Canales, J., Costas, M.J., and Cameselle, J.C. (2014). Bifunctional homodimeric triokinase/FMN cyclase: contribution of protein domains to the activities of the human enzyme and molecular dynamics simulation of domain movements. J. Biol. Chem. *289*, 10620–10636.
3. Tyynismaa, H., Carroll, C.J., Raimundo, N., Ahola-Erkkilä, S., Wenz, T., Ruhanen, H., Guse, K., Hemminki, A., Peltola-Mjøsund, K.E., Tulkki, V., et al. (2010). Mitochondrial myopathy induces a starvation-like response. Hum. Mol. Genet. *19*, 3948–3958.
4. Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum. Mutat. *36*, 928–930.
5. Mayr, J.A., Haack, T.B., Graf, E., Zimmermann, F.A., Wieland, T., Haberberger, B., Superti-Furga, A., Kirschner, J., Steinmann, B., Baumgartner, M.R., et al. (2012). Lack of the mitochondrial protein acylglycerol kinase causes Sengers syndrome. Am. J. Hum. Genet. *90*, 314–320.
6. Garbade, S.F., Shen, N., Himmelreich, N., Haas, D., Trefz, F.K., Hoffmann, G.F., Burgard, P., and Blau, N. (2019). Allelic phenotype values: a model for genotype-based phenotype prediction in phenylketonuria. Genet. Med. *21*, 580–590.
7. Hers, H.G., and Kusaka, T. (1953). The metabolism of fructose-1-phosphate in the liver. Biochim. Biophys. Acta *11*, 427–437.
8. Rodrigues, J.R., Cameselle, J.C., Cabezas, A., and Ribeiro, J.M. (2019). Closure of the Human TKFC Active Site: Comparison of the Apoenzyme and the Complexes Formed with Either Triokinase or FMN Cyclase Substrates. Int. J. Mol. Sci. *20*, 20.
9. Siebold, C., Arnold, I., Garcia-Alles, L.F., Baumann, U., and Erni, B. (2003). Crystal structure of the Citrobacter freundii dihydroxyacetone kinase reveals an eight-stranded α-helical barrel ATP-binding domain. J. Biol. Chem. *278*, 48236–48244.
10. Jang, C., Hui, S., Lu, W., Cowan, A.J., Morscher, R.J., Lee, G., Liu, W., Tesz, G.J., Birnbaum, M.J., and Rabinowitz, J.D. (2018). The Small Intestine Converts Dietary Fructose into Glucose and Organic Acids. Cell Metab. *27*, 351–361.e3.
11. Kinote, A., Faria, J.A., Roman, E.A., Solon, C., Razolli, D.S., Ignacio-Souza, L.M., Sollon, C.S., Nascimento, L.F., de Araújo, T.M., Barbosa, A.P., et al. (2012). Fructose-induced hypothalamic AMPK activation stimulates hepatic PEPCK and gluconeogenesis due to increased corticosterone levels. Endocrinology *153*, 3633–3645.
12. Molin, M., Norbeck, J., and Blomberg, A. (2003). Dihydroxyacetone kinases in Saccharomyces cerevisiae are involved in detoxification of dihydroxyacetone. J. Biol. Chem. *278*, 1415–1423.

13. Steinmann, B.S.R. (2016). Disorders of Fructose Metabolism. In Inborn Metabolic Diseases, B.M, J.M. Saudubray and J. Walter, eds. (Berlin, Heidelberg: Springer), pp. 161–168.
14. Hannou, S.A., Haslam, D.E., McKeown, N.M., and Herman, M.A. (2018). Fructose metabolism and metabolic disease. J. Clin. Invest. *128*, 545–555.
15. Molin, M., Pilon, M., and Blomberg, A. (2007). Dihydroxyacetone-induced death is accompanied by advanced glycation endproduct formation in selected proteins of Saccharomyces cerevisiae and Caenorhabditis elegans. Proteomics *7*, 3764–3774.
16. Franke, S., Dawczynski, J., Strobel, J., Niwa, T., Stahl, P., and Stein, G. (2003). Increased levels of advanced glycation end products in human cataractous lenses. J. Cataract Refract. Surg. *29*, 998–1004.
17. Hashim, Z., and Zarina, S. (2011). Advanced glycation end products in diabetic and non-diabetic human subjects suffering from cataract. Age (Dordr.) *33*, 377–384.
18. Shamsi, F.A., Lin, K., Sady, C., and Nagaraj, R.H. (1998). Methylglyoxal-derived modifications in lens aging and cataract formation. Invest. Ophthalmol. Vis. Sci. *39*, 2355–2364.
19. Smith, K.R., Hayat, F., Andrews, J.F., Migaud, M.E., and Gassman, N.R. (2019). Dihydroxyacetone Exposure Alters NAD(P)H and Induces Mitochondrial Stress and Autophagy in HEK293T Cells. Chem. Res. Toxicol. *32*, 1722–1731.
20. Smith, K.R., Granberry, M., Tan, M.C.B., Daniel, C.L., and Gassman, N.R. (2018). Dihydroxyacetone induces G2/M arrest and apoptotic cell death in A375P melanoma cells. Environ. Toxicol. *33*, 333–342.
21. Marco-Rius, I., von Morze, C., Sriram, R., Cao, P., Chang, G.Y., Milshteyn, E., Bok, R.A., Ohliger, M.A., Pearce, D., Kurhanewicz, J., et al. (2017). Monitoring acute metabolic changes in the liver and kidneys induced by fructose and glucose using hyperpolarized [2-^{13} C]dihydroxyacetone. Magn. Reson. Med. *77*, 65–73.
22. Balasubramaniam, S., Christodoulou, J., and Rahman, S. (2019). Disorders of riboflavin metabolism. J. Inherit. Metab. Dis. *42*, 608–619.
23. Cabezas, A., Costas, M.J., Pinto, R.M., Couto, A., and Cameselle, J.C. (2005). Identification of human and rat FAD-AMP lyase (cyclic FMN forming) as ATP-dependent dihydroxyacetone kinases. Biochem. Biophys. Res. Commun. *338*, 1682–1689.
24. Diao, F., Li, S., Tian, Y., Zhang, M., Xu, L.G., Zhang, Y., Wang, R.P., Chen, D., Zhai, Z., Zhong, B., et al. (2007). Negative regulation of MDA5- but not RIG-I-mediated innate antiviral signaling by the dihydroxyacetone kinase. Proc. Natl. Acad. Sci. USA *104*, 11706–11711.
25. Komuro, A., Bamming, D., and Horvath, C.M. (2008). Negative regulation of cytoplasmic RNA-mediated antiviral signaling. Cytokine *43*, 350–358.
26. Jing, H., and Su, H.C. (2019). New immunodeficiency syndromes that help us understand the IFN-mediated antiviral immune response. Curr. Opin. Pediatr. *31*, 815–820.
27. Rice, G.I., Del Toro Duany, Y., Jenkinson, E.M., Forte, G.M., Anderson, B.H., Ariaudo, G., Bader-Meunier, B., Baildam, E.M., Battini, R., Beresford, M.W., et al. (2014). Gain-of-function mutations in IFIH1 cause a spectrum of human disease phenotypes associated with upregulated type I interferon signaling. Nat. Genet. *46*, 503–509.
28. Asgari, S., Schlapbach, L.J., Anchisi, S., Hammer, C., Bartha, I., Junier, T., Mottet-Osman, G., Posfay-Barbe, K.M., Longchamp, D., Stocker, M., et al. (2017). Severe viral respiratory infections in children with *IFIH1* loss-of-function mutations. Proc. Natl. Acad. Sci. USA *114*, 8342–8347.
29. Lamborn, I.T., Jing, H., Zhang, Y., Drutman, S.B., Abbott, J.K., Munir, S., Bade, S., Murdock, H.M., Santos, C.P., Brock, L.G., et al. (2017). Recurrent rhinovirus infections in a child with inherited MDA5 deficiency. J. Exp. Med. *214*, 1949–1972.
30. Xu, L., Xiao, N., Liu, F., Ren, H., and Gu, J. (2009). Inhibition of RIG-I and MDA5-dependent antiviral response by gC1qR at mitochondria. Proc. Natl. Acad. Sci. USA *106*, 1530–1535.
31. Takashima, K., Oshiumi, H., Takaki, H., Matsumoto, M., and Seya, T. (2015). RIOK3-mediated phosphorylation of MDA5 interferes with its assembly and attenuates the innate immune response. Cell Rep. *11*, 192–200.
32. Broquet, A.H., Hirata, Y., McAllister, C.S., and Kagnoff, M.F. (2011). RIG-I/MDA5/MAVS are required to signal a protective IFN response in rotavirus-infected intestinal epithelium. J. Immunol. *186*, 1618–1626.
33. Feichtinger, R.G., Oláhová, M., Kishita, Y., Garone, C., Kremer, L.S., Yagi, M., Uchiumi, T., Jourdain, A.A., Thompson, K., D'Souza, A.R., et al. (2017). Biallelic C1QBP Mutations Cause Severe Neonatal-, Childhood-, or Later-Onset Cardiomyopathy Associated with Combined Respiratory-Chain Deficiencies. Am. J. Hum. Genet. *101*, 525–538.

The American Journal of Human Genetics, Volume *106*

Supplemental Data

Bi-allelic Variants in *TKFC* Encoding Triokinase/FMN Cyclase Are Associated with Cataracts and Multisystem Disease

Saskia B. Wortmann, Brigitte Meunier, Lamia Mestek-Boukhibar, Florence van den Broek, Elaina M. Maldonado, Emma Clement, Daniel Weghuber, Johannes Spenger, Zdenek Jaros, Fatma Taha, Wyatt W. Yue, Simon J. Heales, James E. Davison, Johannes A. Mayr, and Shamima Rahman

SUPPLEMENTARY INFORMATION

Supplementary Figure 1:

Sequence comparison between human TKFC and yeast DAK1 and DAK2

% similarity human TKFC and DAK1:36%; human TKFC and DAK2:37%

```
Dak1          -MSAKSFEVTD-PVNSSLKGFALANPSITLVPEEKILFRKTD------SDKIALISGGGS  52
Dak2          -MSHKQFKSDGNIVTPYLLGLARSNPGLTVIKHDRVVFRTASAPNSGNPPKVSLVSGGGS  59
human-Dak     MTSKKLVNSVAGCADDALAGLVACNPNLQLLQGHRVALRSDLDS---LKGRVALLSGGGS  57
                * * .:     .   * *:. .**.: ::  .:: :*.          :::*:*****

Dak1          GHEPTHAGFIGKGMLSGAVVGEIFASPSTKQILNAIRLVNE-NASGVLLIVKNYTGDVLH 111
Dak2          GHEPTHAGFVGEGALDAIAAGAIFASPSTKQIYSAIKAVE--SPKGTLIIVKNYTGDIIH 117
human-Dak     GHEPAHAGFIGKGMLTGVIAGAVFTSPAVGSILAAIRAVAQAGTVGTLLIVKNYTGDRLN 117
              ****:****:*:* * .  .* :*:**:. .*  **: *   .  *.*:******** ::

Dak1          FGLSAERARALGINCRVAVIGDDVAVGREKGGMVGRRALAGTVLVHKIVGAFAEEYSSKY 171
Dak2          FGLAAERAKAAGMKVELVAVGDDVSVGKKKGSLVGRRGLGATVLVHKIAGAAASH---GL 174
human-Dak     FGLAREQARAEGIPVEMVVIGDDSAFTVLK--KAGRRGLCGTVLIHKVAGALAEA---GV 172
              ***: *:*:* *:  .:..:*** :.   *   .***.* .***:**:.** *.

Dak1          GLDGTAKVAKIINDNLVTIGSSLDHCKVPGRKFESELNEKQMELGMGIHNEPGVKVLDPI 231
Dak2          ELAEVAEVAQSVVDNSVTIAASLDHCTVPGHKPEAILGENEYEIGMGIHNESGTYKSSPL 234
human-Dak     GLEEIAKQVNVVAKAMGTLGVSLSSCSVPGSKPTFELSADEVELGLGIHGEAGVRRIKMA 232
               *   *: .: : .   *:. **. *.*** *    *. .: *:*:***.* *.   .

Dak1          PSTEDLISKYMLPKLLDPNDKDRAFVKFDEDDEVVLLVNNLGGVSNFVISSITSKTTDFL 291
Dak2          PSISELVSQ-MLPLLL-DEDEDRSYVKFEPKEDVVLMVNNMGGMSNLELGYAAEVISEQL 292
human-Dak     T--ADEIVKLMLDHMT--NTTNASHVPVQPGSSVVMMVNNLGGLSFLELGIIADATVRSL 288
                  : : : **  :   :  : :.* .:  ..**::***:**:* : :.  :.     *

Dak1          KENYNITPVQTIAGTLMTSFNGNGFSITLLNATKATKALQSDFEEIKSVLDLLNAFTNAP 351
Dak2          IDKYQIVPKRTITGAFITALNGPGFGITLMNASKAGGDILKYFDYPTTASGWNQMYHSAK 352
human-Dak     E-GRGVKIARALVGTFMSALEMPGISLTLLLVDEP---LLKLIDAETTAAAWPNVAAV-- 342
                   :   :::.*:::::::  *:.:**: . :    : . ::  .:.    :

Dak1          GWPIADFEKT-SAPSVNDDLLHNEVTAKAVGTYDFDKFAEWMKSGAEQVIKSEPHITELD 410
Dak2          DWEVLAKGQVPTAPSLK--TLRNEK--GSGVKADYDTFAKILLAGIAKINEVEPKVTWYD 408
human-Dak     --SITGRKRSRVAPAEPQEAPDS----TAAGGSASKRMALVLERVCSTLLGLEEHLNALD 396
                 :    :   **:       .     :      . :*  :      :   * ::.  *

Dak1          NQVGDGDCGYTLVAGVKGITENLDKLS--KDSLSQAVAQISDFIEGSMGGTSGGLYSILL 468
Dak2          TIAGDGDCGTTLVSGGEALEEAIKNHTLRLEDAALGIEDIAYMVEDSMGGTSGGLYSIYL 468
human-Dak     RAAGDGDCGTTHSRAARAIQEWLKEGPP-PASPAQLLSKLSVLLLEKMGGSSGALYGLFL 455  G445S
                .****** *   . ..: * :.:      . :  : .:: ::  .***:**.**.: *

Dak1          SGFSHGLIQVCKSKDEPVTKEIVAKSLGIALDTLYKYTKARKGSSTMIDALEPFVKEFTA 528
Dak2          SALAQGVRD---SGDKELTAETFKKASNVALDALYKYTRARPGYRTLIDALQPFVEALKA 525
human-Dak     TAAAQPLKAK-------TSLPAWSAAMDAGLEAMQKYGKAAPGDRTMLDSLWAAGQELQA 508
              :. :: :           :      : . .*::: ** :*  *  *::*:*    : : *

Dak1          SKDFN--------KAVKAAEEGAKSTATFEAKFGRASYVGDS--------SQVEDPGAVG 572
Dak2          GKGPR--------AAAQAAYDGAEKTRKMDALVGRASYVAKEELRKLDSEGGLPDPGAVG 577
human-Dak     WKSPGADLLQVLTKAVKSAEAAAEATKNMEAGAGRASYISSARL-------EQPDPGAVA 561  R543I
               *.           *.::*  .*: * .::*  *****:..             *****.

Dak1          LCEFLKGVQSAL*--        584
Dak2          LAALLDGFVTAAGY*        591
human-Dak     AAAILRAILEVLQS-        575
```

Supplementary Table 1: Results of laboratory investigations

Parameter (reference range)	**Patient 1-1**	**Patient 1-2**	**Patient 2-1**	**Patient 2-2**
Full blood count	**Microcytic anaemia**	Normal	Normal	**Microcytic anaemia**
Vacuolated lymphocytes in peripheral blood film	Not seen	Not seen	ND	ND
Renal Function	Normal	Normal	Normal	Normal
Liver function: Albumin (35-52 g/L) ALT (10-25 U/L) Prothrombin time (8.2-14.1 seconds) INR (<1.2)	Normal 45 29 ND ND	**Abnormal:** **23** **900** 15.1-**23.5** ND	Normal	**Abnormal:** **23 - 38** **20- 386** ND < 1.2 -**1.4**
Blood lactate (<2 mmol/L)	1.2	**16**	Normal	0.9-**2.4**
Ammonia (<40 umol/L)	ND	25	ND	18
FGF21 (44 – 1515 pg/mL)	20	**4350**	ND	ND
Blood carnitine profile	Normal	Normal	ND	Normal
Plasma amino acids (Phenylalanine 21-93 µmol/l)	Normal profile	Normal profile	Mildly elevated phenylalanine	Phenylalanine 376-404
Very long chain fatty acids	Normal profile	Normal profile	ND	ND
Phytanate and pristanate	Normal	Normal	ND	ND
Transferrin electrophoresis	ND	Normal glycoforms	ND	ND
Cholesterol (112-189 mg/dl)	Normal	Normal	ND	Normal 91-147
Triglycerides (29-102 mg/dl)	Normal	Normal	ND	Elevated 58-**604**
Creatine kinase (75-230 U/L)	86	125	ND	ND
Urine organic acids	Elevated pyruvate, 3-hydroxbutyrate and acetoacetate	Mildly raised pyruvate (and dopamine metabolites reflecting inotrope therapy)	ND	3-methylglutaconic aciduria and elevated lactate excretion on one occasion, but normal at other times
Urine reducing substances	Negative	ND	ND	ND
Urine galactitol (3-17 mmol/mol creatinine)	**31**	ND	ND	ND

Urine lactose (0-34 mmol/mol creatinine)	**61**	ND	ND	ND
Other urine polyols	Normal	ND	ND	ND
Urine glycosaminoglycans	Not elevated	ND	ND	ND
Urine/plasma guanidino-acetate and creatine	Normal	ND	ND	ND
Red cell galactose-1-phosphate (<0.10 umol/gHb)	<0.10	ND	ND	ND
Galactose-1-phosphate uridyltransferase (15-35 umol/h/g Hb)	23.3	ND	ND	ND
Galactokinase (0.9-2.2 mol/hr/g Hb)	2.2	ND	ND	ND
Biotinidase (3.9 - 18.9 nmol/mL/min)		5.9	ND	ND
Bloodspot acid maltase activity	ND	Normal	ND	ND
Viral serology	ND	Negative	ND	ND

Key: ALT Alanine aminotransferase, FGF21 fibroblast growth factor 21

Supplementary methods

Whole genome sequencing in Family 1

1. **Genomic DNA (gDNA) preparation:** Whole genome sequencing (WGS) was performed on the trio of father, mother and affected individual P1-2. gDNA was extracted from blood samples in a diagnostic accredited lab (NE Thames Regional Genetics Lab).

   High quality gDNA was used for whole genome library preparation. 1ul gDNA was run on a 1% agarose gel to confirm absence of degradation. gDNA concentration was measured using Cubit dsDNA Broad Range Assay Kit (Invitrogen product #Q32850).

   DNA was diluted to 1.1ug in total volume of 55uL in HT1 buffer and transferred to Covaris 50uL individual tubes. gDNA was sheared to 350bp using E220 Focused-ultrasonicators (Covaris) for 60 seconds with the following parameters: target peak BP 400, Peak Incident Power 140, Duty factor 10% and 22 cycles of burst. Successful shearing was assessed on 1% agarose gel prior to starting library preparation.

2. **Library Preparation:** Whole genome gDNA libraries were prepared using TruSeq DNA PCR-Free Library Prep (Illumina FC-121-3001) following manufacturer advice starting with 1ug of sheared gDNA (in 50uL). Libraries were single indexed using Illumina's indexed adapters (Set A FC-121-3001 or Set B FC-121-3002). Library concentration was measured using quantitative PCR (qPCR) following the manufacturer's advice (KAPA Biosystems). Briefly, 2ul of library was diluted 10000x and 20000x in dilution buffer (100uL Tween, 2mL 1M Tris and 198mL dH_2O) and incubated overnight. qPCR was performed in triplicates in a total volume of 16uL each and run on an Applied Biosystem 7300 qPCR machine.

3. **Library normalization and sequencing:** Libraries were normalised to 2nM with Tris HCl (10mM) pH8.5 supplemented with 0.1% tween 20. Libraries were denatured with 0.2N NaOH and stabilised with 200mMTris HCl. gDNA libraries for the parents were pooled and sequenced on an Illumina NextSeq550 with a 2.7 pM loading concentration. The proband was sequenced on a double flow cell on an Illumina HiSeq Rapid Mode starting with 9pM loading concentration.

4. **Bioinformatics:** Read mapping and variant calling: Basecalling of raw sequencing reads was performed on BaseSpace Sequence hub (basespace.illumina.com). Fastq files for each individual were downloaded from BaseSpace, and reads from different lanes were merged together. Mapping and variant calling were performed using a GENALICE appliance running GENALICE Map 2.5.5 including Mapping, Variant Calling and the Population Calling module for trio analysis (GENALICE BV, Netherlands). GRCh build 37 and GENALICE default configuration files were used for WGS mapping, and trio variant detection. Aligned reads were stored in the GAR format (GENALICE Aligned Reads), using less than 5GB per sample. Variants were stored in a GVM (GENALICE Variant Map) per trio, using less than 200MB per sample. A standard multi sample VCF with Mendelian inheritance annotation using Context Based Call Enhancement was extracted from each GVM.

5. **Variant interpretation:**
   - **i. Pre-filtering step**
     a. Common variants were filtered out (≤ 0.5% in 1000G[1], ExAC[2] and Exome Variant Server (evs.gs.washington.edu/EVS/) databases). For homozygous and hemizygous variants in proband, Minor Allele Frequency (MAF) was increased to ≤10% and variants with no homozygotes/ hemizygotes in ExAC were investigated.

b. Effect of variant on protein function was set to include predicted pathogenic, likely pathogenic and uncertain significance (benign or likely benign variants were investigated if further evidence of pathogenicity was available).
c. Variants associated with loss of function were kept, causing either: frameshift, in-frame in/del, missense or splice site (±7 nucleotides) alterations.
d. 5'UTR and 3'UTR variants were also investigated for genes known to be disease-causing or with compelling evidence for candidate genes.

ii. Exclusion of known genetic causes: This comprised setting a virtual gene panel as a filter to investigate genes associated with the patients' reported phenotypes as the first line of investigation. The gene panel was constructed by converting clinical phenotypes to HPO terms retrieving associated genes from different sources: The *Genomics England* PanelApp (https://panelapp.genomicsengland.co.uk/),Phenotips (https://phenotips.org/) OMIM Gene Map (https://www.omim.org/search/advanced/geneMap), established panels, UCL Great Ormond street Institute of Child Health disease experts and literature search in PubMed. A broader gene panel was analysed which consisted of variants in disease-associated genes from OMIM and DDG2P[3] databases.

6. **Allele Segregation**: allele segregation of the *TKFC* variant (c.1628G/T) was analysed in the family by PCR amplification followed by capillary sequencing (Sanger sequencing) using the following PCR primer set:

TKFC_forward primer: 5'-TCCCTGCTGGAAGTAGATGAG-3'
TKFC_reverse primer: 5'-CTGCAAGACCTCCAAGATGG-3'

Exome sequencing in Family 2

Exome sequencing was performed using the SureSelect Human All Exon 60 Mb Kit (Agilent) for enrichment and a HiSeq4000 (Illumina) for sequencing. The average coverage was 93-fold in P2-1, 138-fold in P2-2, 108-fold in the mother, and 126-fold in the father. The 20-fold coverage was 97.2% in P2-1, 99.0% in P2-2, 98.3% in the mother, and 98.3% in the father. We used BWA (version 0.5.8) for read alignment to the human reference assembly (hg19). Single-nucleotide variants (SNVs) and small insertions and deletions were detected with SAMtools as well as GATK. We excluded variants present with a frequency higher than 1% in 18000 control exomes in our database and public databases including gnomAD. Autosomal recessive, autosomal dominant, X-linked and inheritance via the mitochondrial DNA were considered.

Skin fibroblast culture and western blot in Family 1

Skin fibroblasts were subcultured in DMEM medium supplemented with 10% FBS, 10,000 U/mL penicillin-streptomycin, and 100 µg/mL uridine. Multiple cell passages were used as replicates for TKFC patient fibroblasts and multiple passages of fibroblasts from two metabolic disease control patients were used as controls. All cells were confirmed to be mycoplasma-free and collected at 90% confluency for protein extraction.

Protein was extracted with cOmplete™ Mini EDTA-free Protease Inhibitor Cocktail in RIPA buffer for 30 minutes on ice. The lysed cell solution was centrifuged for 10 minutes at 16,000 xg and protein containing supernatant was collected. Protein concentration was measured by Pierce BCA assay kit as per the manufacturer's instructions.

Loading buffer (x6 solution) was added to samples prior to heat treatment for 10 mins at 70°C. Protein samples (30 µg of protein per lane) and ladder (Precision Plus Protein Kaleidoscope) were loaded into

Novex tris-glycine SDS 10% gels and SDS-PAGE ran at 120V for approximately 60 minutes in tris-glycine SDS running buffer.

Proteins were transferred onto Trans-Blot Turbo PVDF membrane via a Trans-Blot Turbo transfer system (Bio-Rad) at mixed molecular weight setting (1.3 A, 25 V for 7 minutes).

Blots were blocked with 0.1% skimmed milk-tris buffered saline-0.05% Tween 20 solution for 1 hour at room temperature. Primary antibodies were added to blocking solution (St Johns Laboratory Rabbit Anti-TKFC, STJ117616 at 1:1000; Proteintech Rabbit Anti-Vinculin; 26520-1-AP at 1:2000) and blots were probed overnight at 4°C. Membranes were washed for 5 minutes 3x prior to probing with secondary antibody (Cell Signaling Technology Goat Anti-Rabbit, 7074 at 1:2000 in blocking buffer) for 1 hour at room temperature. Blots were washed for 5 minutes x3 prior to visualisation by Thermo Scientific Pierce ECL Western blotting substrate (Vinculin after 5 second exposure; TKFC after 33 second exposure).

ImageJ 1.51j8 Java 1.8.0_112 (64-bit) was used to measure protein band density after converting images to 8-bit images; and GraphPad Prism v8.1.1 (330) for Windows 64-bit was used for graph generation and data analysis.

Preparation of recombinant TKFC protein

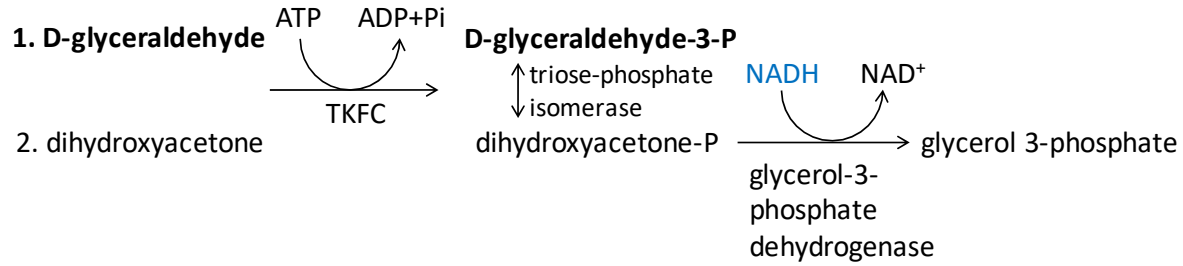
Wild-type human TKFC gene was cloned into the BamHI (…TAAGGATCCg**ATG**ACCTCC…) and BglII (…CAGAGC**TAA**AGATCTGCA…) sites of the pRSET-B vector providing an N-terminal 6xHis-tag. The two variants p.R543I and p.G445S were generated by site-directed mutagenesis (NEB-Builder, New England Biolabs) of the wild-type plasmid. The wild-type and mutated plasmids from several clones were Sanger sequenced in order to verify the mutated positions and to exclude any other variants. Wild-type or recombinant protein was expressed in *Escherichia coli* strain BL21(DE3)pLysS (Promega) grown on LB medium (yeast extract 5 g/l, tryptone 10 g/l, NaCl 5 g/l, pH 7.4) containing ampicillin (100 mg/l) and chloramphenicol (50 mg/l). Overnight cultures were inoculated into 100 ml of growth medium at a cell density of 0.1 optical density at 600 nm (OD600) and grown to an OD600 of 0.3 at 37°C under shaking at 200 rpm (revolutions per minute). The expression of recombinant TKFC was induced by the addition of 0.5 mmol/l IPTG (isopropyl β-D-1-thiogalactopyranoside) and shifting the temperature to 30 °C. The cells were harvested after three hours by centrifugation at 4500 rpm for 5 minutes. Cell pellets were washed with sterile water and resuspended in 6 ml of equilibration buffer (300 mmol/l NaCl, 10 mmol/l imidazole, 50 mmol/l sodium phosphate pH 7.4) and shock-frozen in liquid nitrogen. All following purification steps were performed either under ice-cooling or at 4°C (centrifugation steps).

After thawing, the cell suspensions were treated by ultrasonification with a micro sonification tip and sonification blast for 0.5 seconds and break for 2.5 seconds (Branson digital sonifier). Total sonification time, including brakes, was 5 minutes. After sonification the cell homogenates were centrifuged at 15000 x g for 20 minutes. The supernatants were loaded on HisPur cobalt spin columns (1 ml columns, Pierce Biotechnology)[4]. Spin columns had been washed twice with equilibration buffer before loading by centrifugation (700 x g, 2 minutes) to remove the column storage buffer. After the first 2 ml aliquot of the supernatant of the *E. coli* cell homogenates was loaded, the columns were closed on both ends, incubated for 10 minutes under ice cooling and shaking. Then they were centrifuged to separate the flow-through (700 x g, 2 minutes). This step was repeated since the volume of the cell homogenate exceeded the loading volume of the columns. After completing the loading step, the columns were washed three times with equilibration buffer. Finally, the His-tagged proteins were eluted with 3 x 1

ml of elution buffer (300 mmol/l NaCl, 150 mmol/l imidazole, 50 mmol/l sodium phosphate pH 7.4). The elution was performed by allowing the elution buffer to drip through (no centrifugation) and collecting the eluent in separate tubes. In order to remove the imidazole containing elution buffer, the eluent was diluted with dilution buffer (30 mmol/l NaCl, 30 mmol/l Tris pH 8.0) to a final volume of 15 ml, loaded on ultra centrifugal devices (Amicon Ultra 15, size exclusion 10 kDa) and centrifuged at 2500 x g for 25 minutes to reduce the volume to approximately 500 µl. This washing step was repeated twice and the remaining protein solution was stored in aliquots, which were shock frozen with liquid nitrogen and stored at -80 °C. Recombinant protein was quantified by polyacrylamide electrophoresis and equal amounts of wild-type and variant TKFC protein were used for enzymatic assays. At least two biological replicates were performed for wild-type and mutated proteins.

TKFC enzyme assay

The phosphorylation activity of TKFC was measured with either D-glyceraldehyde (GA) or dihydroxyacetone (DHA) as substrates in coupled assays, following the decrease of NADH at 340 nm by spectrophotometry in a 96 well plate reader at 37 °C for 20 minutes[5]:



The reaction mixture for GA phosphorylation contained a final concentration of 100 mmol/l Tris-HCl, pH 7.5, 0.18 mmol/l NADH, 10 mmol/l $MgCl_2$, 5 mmol/l ATP, 3 units/ml glycerol-3-phosphate dehydrogenase, 15 units/ml triose-phosphate isomerase and 0.1 mg/ml bovine serum albumin. Recombinant TKFC protein (25 µl of either wild-type, p.G455S or p.R543I) was added to the above mentioned mixture and pre-incubated for 5 minutes. The reaction was started by the addition of 12.5 µl of a 10 mmol/l D-glyceraldehyde (GA) stock solution to result in a final concentration of 0.5 mmol/l and final reaction volume of 250 µl.

For measuring the activity on dihydroxyacetone (DHA) the reaction was started by DHA instead of GA and did not contain triose-phosphate isomerase. Each measurement was performed in triplicate and from at least two replicates of recombinant protein purifications. The activity of the mutant proteins was related to wild-type protein, which was set to 100% activity.

Materials and Methods for Yeast Studies

Yeast strains

The control strain BY4742 (*α, his3, leu2, lys2, ura3*) and its derived isogenic deletion strains Δdak1 and Δdak2, were from Euroscarf (Frankfurt, Germany). The double mutant Δdak1Δdak2 was constructed by PCR-based deletion.

Overexpression of yeast DAK2 was obtained by transforming the control strain BY4742 and Δdak mutants with a multi-copy plasmid (yEP352) containing DAK2 under the control of the PGK1 promoter.

Overexpression of human DAK was obtained by transforming the control strain BY4742 and Δdak mutants with a yeast expression plasmid bearing human DAK under the control of the TEF1 promoter (from VectorBuilder, vectorbuilder.com)

Yeast growth media

The pre-culture medium contains 0.7% yeast nitrogen base, 2% glucose, 2% agar and 0.8 g/l of a complete supplement mixture with/without uracil or leucine supplied by Bio 101 (San Diego, CA, USA).

The DHA medium contains 0.7% yeast nitrogen base, 50 mM dihydroxyacetone (DHA), and 0.8 g/l of a complete supplement mixture with/without uracil supplied by Bio 101 (San Diego, CA, USA).

Supplementary References

1. Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., and Abecasis, G.R. (2015). A global reference for human genetic variation. Nature 526, 68-74.
2. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature 536, 285-291.
3. Firth, H.V., Richards, S.M., Bevan, A.P., Clayton, S., Corpas, M., Rajan, D., Van Vooren, S., Moreau, Y., Pettett, R.M., and Carter, N.P. (2009). DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. American journal of human genetics 84, 524-533.
4. Banka, S., de Goede, C., Yue, W.W., Morris, A.A., von Bremen, B., Chandler, K.E., Feichtinger, R.G., Hart, C., Khan, N., Lunzer, V., et al. (2014). Expanding the clinical and molecular spectrum of thiamine pyrophosphokinase deficiency: a treatable neurological disorder caused by TPK1 mutations. Mol Genet Metab 113, 301-306.
5. Rodrigues, J.R., Couto, A., Cabezas, A., Pinto, R.M., Ribeiro, J.M., Canales, J., Costas, M.J., and Cameselle, J.C. (2014). Bifunctional homodimeric triokinase/FMN cyclase: contribution of protein domains to the activities of the human enzyme and molecular dynamics simulation of domain movements. The Journal of biological chemistry 289, 10620-10636.