REPORT

Biallelic Mutations of Methionyl-tRNA Synthetase Cause a Specific Type of Pulmonary Alveolar Proteinosis Prevalent on Réunion Island

Alice Hadchouel,[1,2,3,4,15] Thomas Wieland,[5,15] Matthias Griese,[6,15] Enrico Baruffini,[7,15] Bettina Lorenz-Depiereux,[5] Laurent Enaud,[8] Elisabeth Graf,[5] Jean Christophe Dubus,[9] Sonia Halioui-Louhaichi,[10] Aurore Coulomb,[11] Christophe Delacourt,[1,4,12] Gertrud Eckstein,[5] Ralf Zarbock,[6] Thomas Schwarzmayr,[5] François Cartault,[13] Thomas Meitinger,[5,14] Tiziana Lodi,[7] Jacques de Blic,[1,12] and Tim M. Strom[5,14,*]

Methionyl-tRNA synthetase (MARS) catalyzes the ligation of methionine to tRNA and is critical for protein biosynthesis. We identified biallelic missense mutations in *MARS* in a specific form of pediatric pulmonary alveolar proteinosis (PAP), a severe lung disorder that is prevalent on the island of Réunion and the molecular basis of which is unresolved. Mutations were found in 26 individuals from Réunion and nearby islands and in two families from other countries. Functional consequences of the mutated alleles were assessed by growth of wild-type and mutant strains and methionine-incorporation assays in yeast. Enzyme activity was attenuated in a liquid medium without methionine but could be restored by methionine supplementation. In summary, identification of a founder mutation in *MARS* led to the molecular definition of a specific type of PAP and will enable carrier screening in the affected community and possibly open new treatment opportunities.

Pulmonary alveolar proteinosis (PAP) is characterized by an accumulation of lipoproteins in the pulmonary alveoli; this accumulation leads to restrictive lung disease and respiratory failure.[1–3] PAP is either acquired or inherited in an autosomal-recessive mode. The acquired form (MIM: 610910) affects adults and is attributed to granulocyte-macrophage colony-stimulating factor (GM-CSF) autoantibodies.[4,5] Inherited PAP is usually diagnosed in early childhood. So far, rare mutations in CSF receptor genes *CSF2RA* (MIM: 306250) and *CSF2RB* (MIM: 138981) have been reported as a cause of inherited forms (MIM: 300770, 614370). [6–8] A specific, severe childhood form of PAP is prevalent on Réunion Island, where the incidence is at least 1 in 10,000 newborns.[1,2] Mutations in *CSF2RA* and *CSF2RB* have been excluded previously.[1] Since 1970, approximately 34 children have been diagnosed and treated. If a founder mutation is assumed, the most recent common ancestor of these children can be traced back to the early 18th century.[1] The main symptom is respiratory insufficiency, often leading to death in childhood or adolescence as a result of lung fibrosis despite supportive treatment, including regular whole-lung lavages (Table S1). In addition to lung fibrosis, non-life-threatening liver involvement might be present, as indicated by elevated enzymes, steatosis, fibrosis, or cirrhosis. We investigated 26 DNA samples from individuals who were from Réunion or the nearby islands of Comoros and Madagascar and who were affected with unexplained PAP, and we performed homozygosity mapping and exome and whole-genome sequencing to identify the genetic basis of this disease (Table S2). In addition, we analyzed DNA from a Tunisian sibling pair and from an individual who has sporadic PAP and is living in Paris.[9] Written informed consent was obtained from all study participants. The study was approved by the Comité de Protection des Personnes Île de France II ethical review board.

We performed SNP-array genotyping in 14 affected individuals from the Réunion and Comoros islands by using HumanOmni2.5-4 v.1 and CNV370-Duo v.1 SNP arrays (Illumina); this enabled us to map a homozygous region to chromosome 12q13.3 in all investigated individuals. This region comprised 530 kb between markers rs703817 and rs2277324 (Figure S1) and contained 20 genes. Results from additional SNP genotyping in the two affected siblings from Tunisia were compatible with these findings. Both siblings were homozygous within the critical region on 12q13.3; however, they carried a haplotype different from that found in all affected individuals from Réunion.

[1]Service de Pneumologie et d'Allergologie Pédiatriques, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris, 75015 Paris, France; [2]Université Paris Descartes-Sorbonne Paris Cité, Institut Imagine, 75006 Paris, France; [3]INSERM U-1163, Team of Embryology and Genetics of Congenital Malformations, 75015 Paris, France; [4]INSERM U-955, équipe 4, 94000 Créteil, France; [5]Institute of Human Genetics, Helmholtz Zentrum München, 85764 Neuherberg, Germany; [6]Hauner Children's University Hospital, Ludwig-Maximilians-Universität, German Center for Lung Research, 80337 Munich, Germany; [7]Department of Life Sciences, University of Parma, 43124 Parma, Italy; [8]Department of Pediatrics, Centre Hospitalier Universitaire de La Reunion, 97410 Saint-Pierre, La Reunion, France; [9]Department of Pediatrics, Timone University Hospital, 13385 Marseille, France; [10]Department of Pediatrics, Mongi Slim Hospital La Marsa, University of Tunis El Manar, 1007 Tunis, Tunisia; [11]Service d'Anatomie et Cytologie Pathologiques, Hôpital d'Enfants Armand-Trousseau, Centre Hospitalier Universitaire de Paris-Est, Assistance Publique-Hôpitaux de Paris, 75012 Paris, France; [12]Université Paris-Descartes, 75006 Paris, France; [13]Service de Génétique, Centre Hospitalier Universitaire de la Réunion, 97400 Saint-Denis, La Réunion, France; [14]Institute of Human Genetics, Technische Universität München, 81675 Munich, Germany
[15]These authors contributed equally to this work
*Correspondence: timstrom@helmholtz-muenchen.de
http://dx.doi.org/10.1016/j.ajhg.2015.03.010.

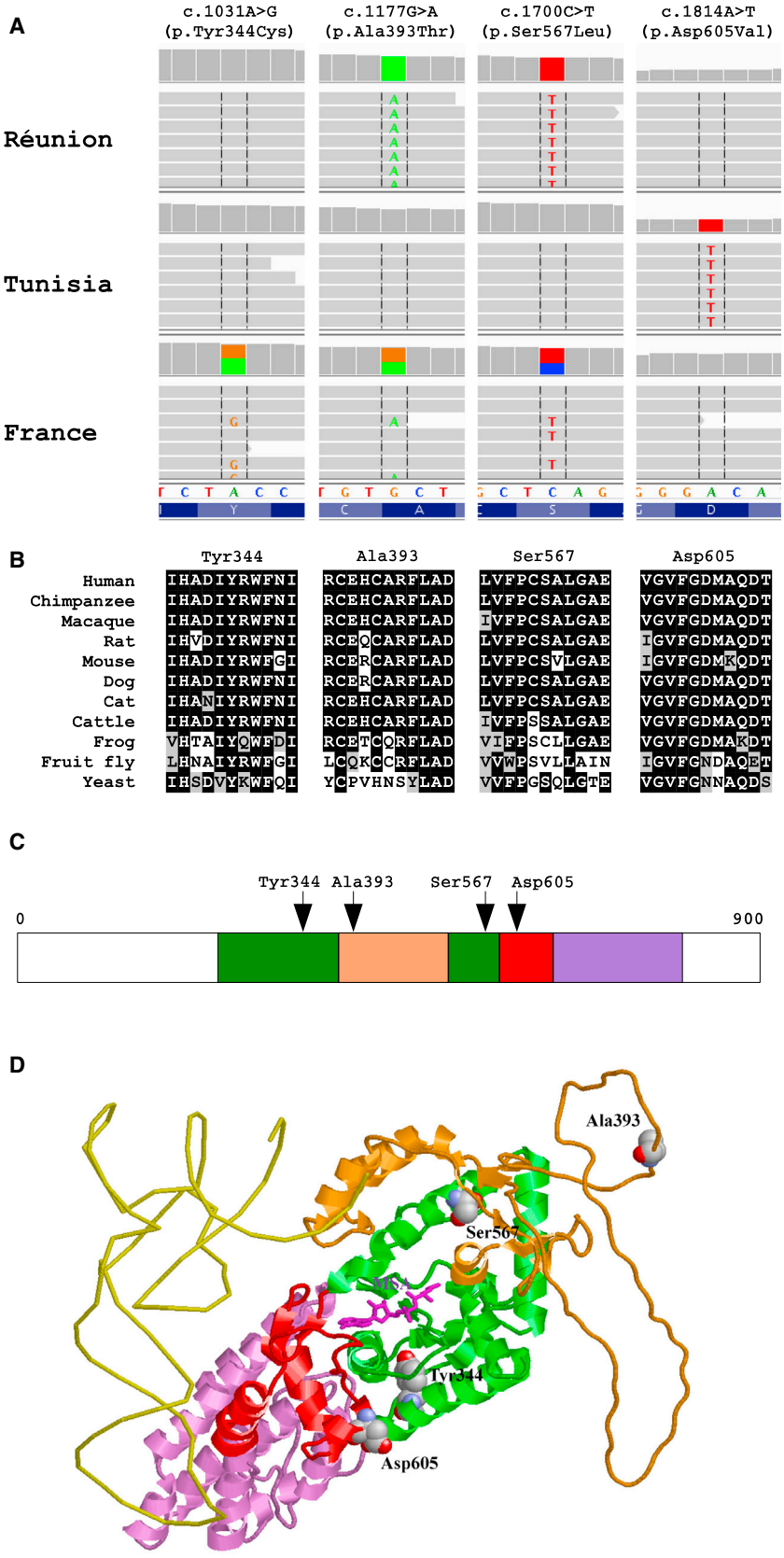


Figure 1. ***MARS*** **Variants in PAP**
(A) Sequencing reads showing the different biallelic variants identified in individuals from Réunion, Tunisia, and France.
(B) Amino acid conservation across MARS orthologs.
(C) Scheme of the domain structure of MARS with the location of the variants.
(D) Predicted tertiary structure. MARS contains a nucleotide-binding (Rossmann) fold (green); a region called the connective polypeptide, which contains the zinc-binding sites (orange); the stem-contact fold domain (red); and the α-helix bundle domain that forms the anticodon-binding site (violet). The positions of the variants are indicated relative to the reference sequence (GenBank: NM_004990.3). The structure of human MARS was predicted by homology modeling based on the *Aquifex aeolicus* structure of MARS complexed with methionyl sulfamoyl adenosine (MSA) and the elongator $tRNA^{Met}$ (PDB: 2CT8)[10] as templates. The model was constructed with the SWISS-MODEL automated protein-structure homology-modeling server. The predicted structure was superimposed with MSA and $tRNA^{Met}$ with SPDBviewer and was visualized with Rasmol.

We performed exome sequencing on two individuals from Réunion and one of the Tunisian siblings to identify possible disease-causing variants. Sequencing was performed via 100-bp paired-end reads on HiSeq2500 systems (Illumina). We generated, on average, 11.3 Gb of sequence, resulting in an average depth of coverage of 135 and in 94% of the target regions' being covered at least 20 times. Single-nucleotide variants (SNVs) and small insertions and deletions were called with SAMtools and Pindel and filtered so that only those variants with a minor-allele frequency (MAF) of less than 1% remained. As controls for filtering, we used 4,000 in-house exomes from individuals with unrelated diseases (Figure S2), the 1000 Genomes Project data (n = 1,700), and the Exome Aggregation Consortium (ExAC) Browser dataset (n = 60,706).

Rare variants common to all three affected individuals were only detected in methionyl-tRNA synthetase (*MARS* [MIM: 156560]), one of the 20 genes in the critical region. Three additional missense variants in the critical region had an allele frequency of at least 0.21 (Table S3). The individuals from Réunion, Comoros, and Madagascar islands carried two homozygous missense variants (GenBank: NM_004990.3): c.1177G>A (p.Ala393Thr; rs141340466) and c.1700C>T (p.Ser567Leu; rs143592405), in exons 10 and 14, respectively. The sibling pair from Tunisia carried

Table 1. Variants Identified in *MARS*

Origin of Individuals	Zygosity	Genome	cDNA	Protein	PPH2	SIFT	CADD
Réunion or Comoros	homozygous	chr12: g.57906083C>T	c.1700C>T	p.Ser567Leu	benign	0.68	17.38
		chr12: g.57894189G>A	c.1177G>A	p.Ala393Thr	benign	0.16	17.63
Tunisia	homozygous	chr12: g.57906594A>T	c.1814A>T	p.Asp605Val	probably damaging	0	26.2
France	heterozygous	chr12: g.57906083C>T	c.1700C>T	p.Ser567Leu	benign	0.68	17.38
		chr12: g.57894189G>A	c.1177G>A	p.Ala393Thr	benign	0.16	17.63
	heterozygous	chr12: g.57892346A>G	c.1031A>G	p.Tyr344Cys	probably damaging	0	23.9

The human genome assembly hg19 (CRCh37) and transcript NM_004990.3 were used as reference sequences. SIFT values below 0.05 are predicted to have functional impact. For CADD, phred-like scores (scaled C scores) are listed.

a different homozygous missense variant, c.1814A>T (p.Asp605Val), in exon 15 (Figure 1, Table 1, and Table S2). The results from Sanger and/or exome sequencing in the remaining 12 affected individuals used for homozygosity mapping and in 12 additional affected individuals from Réunion were consistent with the initial findings. A variant, p.Asp605Gly, was present in a heterozygous state in a single sample of East Asian origin out of the approximately 65,000 samples used for filtering. The variants p.Ala393Thr and p.Ser567Leu carried by the individuals from Réunion were found in four exomes of African origin in the ExAC dataset. In addition, the variant p.Ala393Thr was present in 19 of 4,327 exomes of East Asian origin and in a single exome of different origin. The frequency of the variants p.Ala393Thr and p.Ser567Leu in 1,000 control subjects from Réunion was 22 in 2,000 alleles, resulting in a predicted disease frequency of 1 in 8,264, which is consistent with the observed disease frequency on Réunion. Control subjects were randomly chosen from the Réunion DNA bank, which contains DNA from individuals referred to the Centre Hospitalier Universitaire de la Réunion because of various diagnoses. Individuals with lung diseases were excluded.

Searching for additional disease-causing variants, we sequenced an additional affected individual living in Paris. SNP genotyping was consistent with a constellation in which this individual carried two different haplotypes in the critical region and in which one of these haplotypes was identical with the Réunion haplotype. Indeed, we detected the two variants of the Réunion haplotype and an additional heterozygous non-synonymous substitution, c.1031A>G (p.Tyr344Cys), in exon 9 (Figure 1, Table 1); this substitution was present in a heterozygous state in two exomes of European origin in the ExAC dataset. Capillary sequencing in the parents demonstrated a compound-heterozygous state of both alleles. The Réunion allele was inherited from the mother, who turned out to have been born in Réunion, and the other allele came from the father. We further excluded structural variations at the locus by performing whole-genome sequencing in two affected individuals from Réunion. We generated, on average, 115.5 Gb of sequence, resulting in an average depth of coverage of 32 and in 89% of the coding regions' of the RefSeq collection being covered at least 20 times. In addition, this analysis revealed no evidence for a deletion of any of the coding exons in the critical region.

Homozygous or compound-heterozygous occurrence of rare *MARS* variants was only infrequently observed in control individuals. Our in-house exomes contained only one additional exome carrying a homozygous *MARS* missense variant, c.2180G>A (p.Arg727Gln; rs113808165; MAF = 0.55%), and one compound-heterozygous carrier of the same variant (rs113808165) in combination with the missense variant c.617C>T (p.Pro206Leu; rs138776588; MAF = 0.5%). The two individuals carrying these variants were diagnosed with myocardial infarction and ventricular arrhythmia, respectively. Both missense variants were predicted to be benign by PolyPhen-2 and SIFT. A conservative test comparing the three homozygous or compound-heterozygous variants in the affected individuals with the two homozygous variants found in 4,000 control subjects found the differences to be highly significant (Fisher's exact test: $p < 1.9e{-}8$).

MARS codes for the methionyl-tRNA synthetase, which belongs to the class 1 family of aminoacyl-tRNA synthetases (ARSs). These enzymes play a critical role in protein biosynthesis by charging tRNAs with their cognate amino acids. MARS is a component of a multi-protein complex and catalyzes the ligation of methionine to tRNA molecules. The protein is highly conserved and ubiquitously expressed. Structural prediction of human methionyl-tRNA synthetase (Figure 1 and Figures S3 and S4) showed that Tyr344 and Ser567 lie in the MARS Rossman fold, a domain that contains most of the sites that catalyze both the methionyl adenylation from L-methionine and ATP and the methionylation of the $tRNA^{Met}$. Ala393 lies in a loop in the connective polypeptide, downstream of the first of four conserved CXX[C,D,H] motifs that are involved in the binding of two zinc ions. Asp605 lies in the stem-contact fold domain, which contains both a region that binds to the inside of the L-shaped tRNA and sites that catalyze the methionyl adenylation.[11–13]

Because of the fundamental role of ARSs in cell metabolism, the identified variants most likely result in reduced enzyme activity rather than a complete loss of that activity. Taking advantage of the conservation of MARS between humans and yeast, we assessed enzyme activity of the

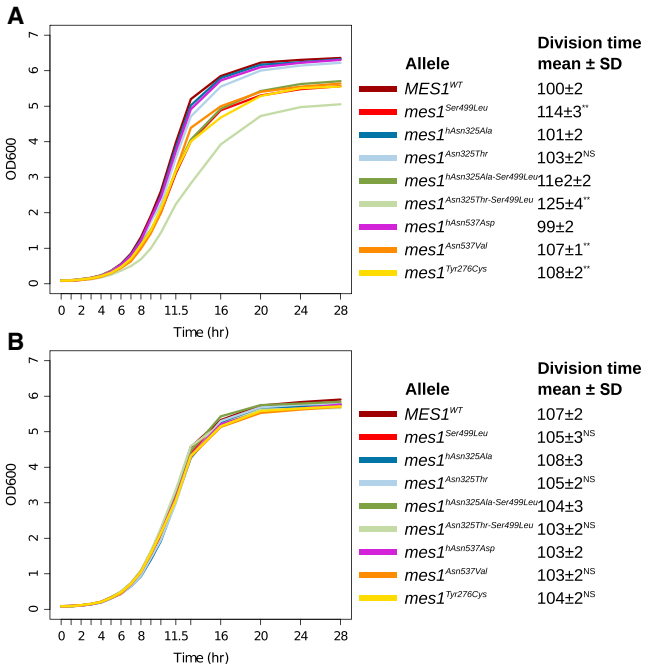


Figure 2. Growth of *MES1* Wild-Type and *mes1* Mutant Strains
Growth (A) without methionine or (B) with 20 μg/ml methionine. Cells were inoculated at the concentration of 0.1 OD_{600}/ml and grown until the stationary phase was reached after 28 hr. At regular intervals, aliquots were used for measurement of cell density by UV-visible spectrophotometry at 600 nm. Sampling times are indicated by x-axis ticks. Tables show division times (minutes) calculated during the exponential phase of growth. Division times are the mean of three independent growth curves. The *S. cerevisiae* strain used in this work was W303-1B (*Matα ade2-1 leu2-3,112 ura3-1 trp1-1 his3-11,15 can1-100*). The *MES1* wild-type allele was cloned in the centromeric vector pFL38[14]. Genomic *MES1* was disrupted in the pFL38*MES1*-transformed W303-1B strain by one-step gene disruption with a KanMX expression cassette.[15] *mes1* mutant and double-mutant alleles were constructed via site-directed mutagenesis through the PCR overlap extension technique with the oligonucleotides listed in Table S4,[16] cloned into vector pFL39[14], and introduced into W303-1B *mes1Δ* pFL38*MES1*. In a second step, strains devoid of pFL38$MES1^{WT}$ and containing the pFL39-borne $MES1^{WT}$ or *mes1* mutant alleles were selected through plasmid shuffling.[17] NS, not significant in a two-tailed, unpaired t test; **$p < 0.01$.

mutated alleles by both growth of wild-type and mutant strains and methionine-incorporation assays in yeast by expressing the variants in the yeast ortholog *MES1* (Figures 2A and 3A and Figures S5 and S6). Activities of the mutated alleles were compared with the wild-type and the humanized alleles in case the amino acid was not conserved between human and yeast (Table 2). Compared with those in wild-type yeast, enzyme activities in humanized alleles were not significantly different. However, compared to those of the wild-type or respective humanized alleles, both calculated division times and ^{35}S incorporation were significantly different in yeast transfected with the mutated alleles, with the exception of $mes1^{Asn325Thr}$, one of the two variants found in the individuals from Réunion. Although $mes1^{Asn325Thr}$ alone does not attenuate enzyme activity, it worsens the phenotype if it is expressed in combination with $mes1^{Ser499Leu}$. Of note, attenuation of enzyme activity occurs in a liquid medium without methionine, and its activity can be restored in a medium supplemented with 20 μg/ml methionine (Figures 2B and 3B). This observation, together with the predicted position of the mutated amino acids inside the protein, renders interference of the variants with substrate binding the most likely functional mechanism.

We have provided convincing genetic and functional evidence that *MARS* mutations are the cause of a specific type of PAP. We delineated a 530-kb candidate region by performing homozygosity mapping in 14 affected individuals from Réunion. Within this candidate region, exome and genome sequences helped to identify two rare homozygous missense variants in close proximity in a single gene, *MARS*. Three other non-synonymous variants in this region had an allele frequency of at least 0.21 and are therefore unlikely to be disease causing. The most likely disease-causing variant, p.Ser567Leu, was found in all 26 investigated affected individuals, had a frequency of only 4 in approximately 60,000 control samples in the ExAC dataset, and was not present in 4,000 in-house control samples. We further excluded structural variations in this region by genome sequencing. Next, we identified a different homozygous *MARS* missense variant in two affected siblings from Tunisia and compound-heterozygous variants, one of which was identical to the Réunion haplotype, in a French individual with sporadic PAP. The variants p.Ala393Thr and p.Ser567Leu are in strong linkage disequilibrium in the Réunion population, therefore hampering a conclusion about their causality for PAP. Functional investigation in yeast provided evidence that p.Ser567Leu is disease causing given that the corresponding mutation in yeast resulted in reduced growth and more than a 50% reduction of methionine incorporation, whereas the mutation corresponding to p.Ala393Thr did not show a phenotype. This interpretation is further supported by data from the East Asian population where the p.Ala393Thr variant is present in 19 of 8,654 alleles, whereas p.Ser567Leu is absent. However, the possibility that p.Ala393Thr contributes to the phenotype cannot be excluded, given that the corresponding mutation in yeast aggravated the phenotype of the mutation corresponding to p.Ser567Leu in a double-mutation strain. We assessed the functional impact of the mutations by conducting yeast complementation studies. Confirmation by aminoacylation assays might be worthwhile. However, the validity of yeast assays has been shown during the investigation of several other ARS mutations in which an attenuated function demonstrated in yeast was confirmed by aminoacylation assays.[18] The human genome harbors 37 ARS loci. 17 encode a cytoplasmic enzyme, 17 a mitochondrial enzyme, and 3 a bi-functional enzyme that is present in both cell compartments.[18] Thus far, ARS mutations have been implicated in autosomal-recessive mitochondrial disease and autosomal-dominant peripheral neuropathies known as Charcot-Marie-Tooth disease.[18]

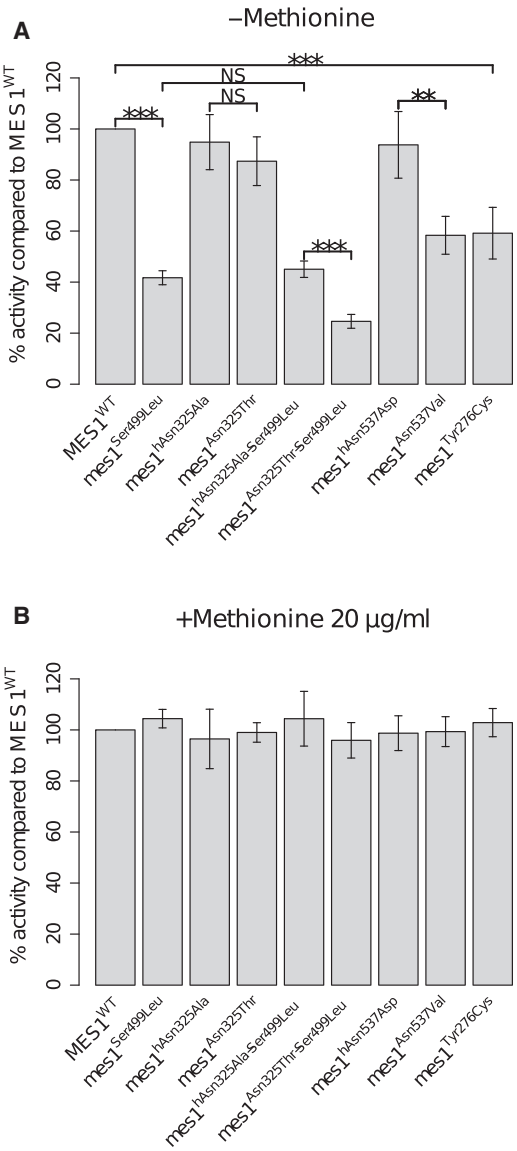


Figure 3. ^{35}S Incorporation of *MES1* Wild-Type and *mes1* Mutant Strains

Incorporation (A) without methionine or (B) with 20 µg/ml methionine. Values are normalized to the wild-type strain, which was set as 100%. Four replicates were performed for the experiments without methionine and three replicates for the experiments with methionine. The error bars indicate SDs. Cells were inoculated at a final concentration of 0.1 OD_{600}/ml in synthetic-complete-dextrose medium (0.69% yeast nitrogen base, 0.1% yeast amino acid and nucleobase mixture, 2% glucose) with or without 20 µg/ml methionine and grown at 37°C. After 16 hr, cells were diluted to a final concentration of 1.2 OD_{600}/ml. After 5 min, cells were supplemented with 1 µl, if grown without methionine, or 10 µl, if grown with methionine, of EasyTag [^{35}S]-protein labeling mix having a specific activity of 1,000 Ci/mmol (Perkin Elmer). Once we verified that the incorporation signal was linear between 2 and 10 min, we blocked protein synthesis after 6 min by adding a mix containing 200 µg cycloheximide, 1 mg erythromycin, 100 µg cold L-methionine, and 100 µg cold L-cysteine and chilling the mixture on ice. We used the trichloroacetic (TCA) method to precipitate total proteins by chilling the cells supplemented with 25% TCA on ice, then resuspended the proteins in 30 µl of 60-mM Tris-HCl (pH 6.8). For each sample, counts per minute/OD_{600} were measured on 10 µl aliquots and normalized to values for the wild-type strain, which was set as 100%. NS, not significant in a two-tailed, paired t test; **$p < 0.01$; ***$p < 0.001$.

Table 2. Variants Introduced into Mes1 for the Studies in Yeast

MARS Variant	Mes1 Variant	Humanized Mes1
p.Ala393Thr	p.Asn325Thr	p.Asn325Ala
p.Ser567Leu	p.Ser499Leu	NA
p.Asp605Val	p.Asn537Val	p.Asn537Asp
p.Tyr344Cys	p.Tyr276Cys	NA

Human transcript NM_004990.3 and yeast protein P00958 were used as reference sequences. The following abbreviation is used: NA, not applicable.

Here, we add a further phenotype to the wide spectrum of tissue-specific human diseases caused by ARS mutations. A single female infant has previously been reported to carry compound-heterozygous *MARS* mutations (c.1108T>C [p.Phe370Leu] and c.1568T>C [p.Ile523Thr]).[19] The described phenotype is compatible with that observed in this study; however, it seems to include additional signs such as acidosis, aminoaciduria, hypothyroidism, and anemia (MIM: 615486). In addition, rare heterozygous *MARS* variants have been reported in individuals affected by late-onset Charcot-Marie-Tooth disease.[20,21] In this study, however, family history of index patients in Réunion does not provide any evidence that heterozygous carriers are affected by Charcot-Marie-Tooth disease. Studies in yeast have demonstrated that attenuation of MARS activity could be rescued by supplementation with methionine. To our knowledge, a similar effect has only been described for a *VARS2* mutation before.[22] This observation suggests the need for investigation of potential beneficial effects of high-dose methionine treatment in humans.

Accession Numbers

The ClinVar accession numbers for the three sequence variants reported in this paper are SCV000196708, SCV000196709, and SCV000196710.

Supplemental Data

Supplemental Data include six figures and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2015.03.010.

Acknowledgments

We thank the families for participating in this study, Sandy Lösecke, Traudl Wesselack, and Andrea Schams for technical assistance, and Abdourahim Chamouine for his support in recruiting patients from Comoros Island. We thank the Centre de Ressources Biologiques de la Réunion (CRB-CHU REUNION UF1646) for providing DNA of controls. This work was supported by the German Ministry of Education and Research (01GM1113B), Else-Kroener Stiftung (MG 2013_A72), the eRARE Project 2009

(EUPAPNet), the European Register and Biobank on Childhood Interstitial Lung Diseases (European Commission, FP7, GA 305653), Telethon (GGP11011), and l'Agence Nationale de la Recherche (ANR-12-BSV1-0004-01).

Received: November 23, 2014
Accepted: March 19, 2015
Published: April 23, 2015

Web Resources

The URLs for data presented herein are as follows:

1000 Genomes, http://browser.1000genomes.org
ClinVar, https://www.ncbi.nlm.nih.gov/clinvar/
ExAC Browser, http://exac.broadinstitute.org
OMIM, http://www.omim.org/
SWISS-MODEL automated protein-structure homology-modeling server, http://swissmodel.expasy.org
UCSC Genome Bioinformatics, http://genome.ucsc.edu

References

1. Enaud, L., Hadchouel, A., Coulomb, A., Berteloot, L., Lacaille, F., Boccon-Gibod, L., Boulay, V., Darcel, F., Griese, M., Linard, M., et al. (2014). Pulmonary alveolar proteinosis in children on La Réunion Island: a new inherited disorder? Orphanet J. Rare Dis. *9*, 85.
2. de Blic, J. (2004). Pulmonary alveolar proteinosis in children. Paediatr. Respir. Rev. *5*, 316–322.
3. Mahut, B., Delacourt, C., Scheinmann, P., de Blic, J., Mani, T.M., Fournet, J.C., and Bellon, G. (1996). Pulmonary alveolar proteinosis: experience with eight pediatric cases and a review. Pediatrics *97*, 117–122.
4. Uchida, K., Beck, D.C., Yamamoto, T., Berclaz, P.Y., Abe, S., Staudt, M.K., Carey, B.C., Filippi, M.D., Wert, S.E., Denson, L.A., et al. (2007). GM-CSF autoantibodies and neutrophil dysfunction in pulmonary alveolar proteinosis. N. Engl. J. Med. *356*, 567–579.
5. Stanley, E., Lieschke, G.J., Grail, D., Metcalf, D., Hodgson, G., Gall, J.A., Maher, D.W., Cebon, J., Sinickas, V., and Dunn, A.R. (1994). Granulocyte/macrophage colony-stimulating factor-deficient mice show no major perturbation of hematopoiesis but develop a characteristic pulmonary pathology. Proc. Natl. Acad. Sci. USA *91*, 5592–5596.
6. Dirksen, U., Nishinakamura, R., Groneck, P., Hattenhorst, U., Nogee, L., Murray, R., and Burdach, S. (1997). Human pulmonary alveolar proteinosis associated with a defect in GM-CSF/IL-3/IL-5 receptor common beta chain expression. J. Clin. Invest. *100*, 2211–2217.
7. Trapnell, B.C., Whitsett, J.A., and Nakata, K. (2003). Pulmonary alveolar proteinosis. N. Engl. J. Med. *349*, 2527–2539.
8. Martinez-Moczygemba, M., Doan, M.L., Elidemir, O., Fan, L.L., Cheung, S.W., Lei, J.T., Moore, J.P., Tavana, G., Lewis, L.R., Zhu, Y., et al. (2008). Pulmonary alveolar proteinosis caused by deletion of the GM-CSFRalpha gene in the X chromosome pseudoautosomal region 1. J. Exp. Med. *205*, 2711–2716.
9. Halioui-Louhaïchi, S., Ben Hariz, M., Hamzaoui, A., Mestiri, T., Jerbi, E., Louzir, B., Ben Khelifa, S., Ben Jaballah, N., De Blic, J., Ben Ammar, M.S., et al. (2005). [Pulmonary alveolar proteinosis in two siblings. A case report]. Tunis. Med. *83*, 488–491.
10. Nakanishi, K., Ogiso, Y., Nakama, T., Fukai, S., and Nureki, O. (2005). Structural basis for anticodon recognition by methionyl-tRNA synthetase. Nat. Struct. Mol. Biol. *12*, 931–932.
11. Mechulam, Y., Schmitt, E., Maveyraud, L., Zelwer, C., Nureki, O., Yokoyama, S., Konno, M., and Blanquet, S. (1999). Crystal structure of Escherichia coli methionyl-tRNA synthetase highlights species-specific features. J. Mol. Biol. *294*, 1287–1297.
12. Schmitt, E., Meinnel, T., Blanquet, S., and Mechulam, Y. (1994). Methionyl-tRNA synthetase needs an intact and mobile 332KMSKS336 motif in catalysis of methionyl adenylate formation. J. Mol. Biol. *242*, 566–576.
13. Casina, V.C., Lobashevsky, A.A., McKinney, W.E., Brown, C.L., and Alexander, R.W. (2011). Role for a conserved structural motif in assembly of a class I aminoacyl-tRNA synthetase active site. Biochemistry *50*, 763–769.
14. Bonneaud, N., Ozier-Kalogeropoulos, O., Li, G.Y., Labouesse, M., Minvielle-Sebastia, L., and Lacroute, F. (1991). A family of low and high copy replicative, integrative and single-stranded S. cerevisiae/E. coli shuttle vectors. Yeast *7*, 609–615.
15. Wach, A., Brachat, A., Pöhlmann, R., and Philippsen, P. (1994). New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. Yeast *10*, 1793–1808.
16. Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., and Pease, L.R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene *77*, 51–59.
17. Baruffini, E., Ferrero, I., and Foury, F. (2010). In vivo analysis of mtDNA replication defects in yeast. Methods *51*, 426–436.
18. Wallen, R.C., and Antonellis, A. (2013). To charge or not to charge: mechanistic insights into neuropathy-associated tRNA synthetase mutations. Curr. Opin. Genet. Dev. *23*, 302–309.
19. van Meel, E., Wegner, D.J., Cliften, P., Willing, M.C., White, F.V., Kornfeld, S., and Cole, F.S. (2013). Rare recessive loss-of-function methionyl-tRNA synthetase mutations presenting as a multi-organ phenotype. BMC Med. Genet. *14*, 106.
20. Gonzalez, M., McLaughlin, H., Houlden, H., Guo, M., Yo-Tsen, L., Hadjivassilious, M., Speziani, F., Yang, X.L., Antonellis, A., Reilly, M.M., and Züchner, S.; Inherited Neuropathy Consortium (2013). Exome sequencing identifies a significant variant in methionyl-tRNA synthetase (MARS) in a family with late-onset CMT2. J. Neurol. Neurosurg. Psychiatry *84*, 1247–1249.
21. Novarino, G., Fenstermaker, A.G., Zaki, M.S., Hofree, M., Silhavy, J.L., Heiberg, A.D., Abdellateef, M., Rosti, B., Scott, E., Mansour, L., et al. (2014). Exome sequencing links corticospinal motor neuron disease to common neurodegenerative disorders. Science *343*, 506–511.
22. Diodato, D., Melchionda, L., Haack, T.B., Dallabona, C., Baruffini, E., Donnini, C., Granata, T., Ragona, F., Balestri, P., Margollicci, M., et al. (2014). VARS2 and TARS2 mutations in patients with mitochondrial encephalomyopathies. Hum. Mutat. *35*, 983–989.

The American Journal of Human Genetics

Supplemental Data

Biallelic Mutations of the Methionyl-tRNA Synthetase Cause a Specific Type of Pulmonary Alveolar Proteinosis Prevalent on Réunion Island

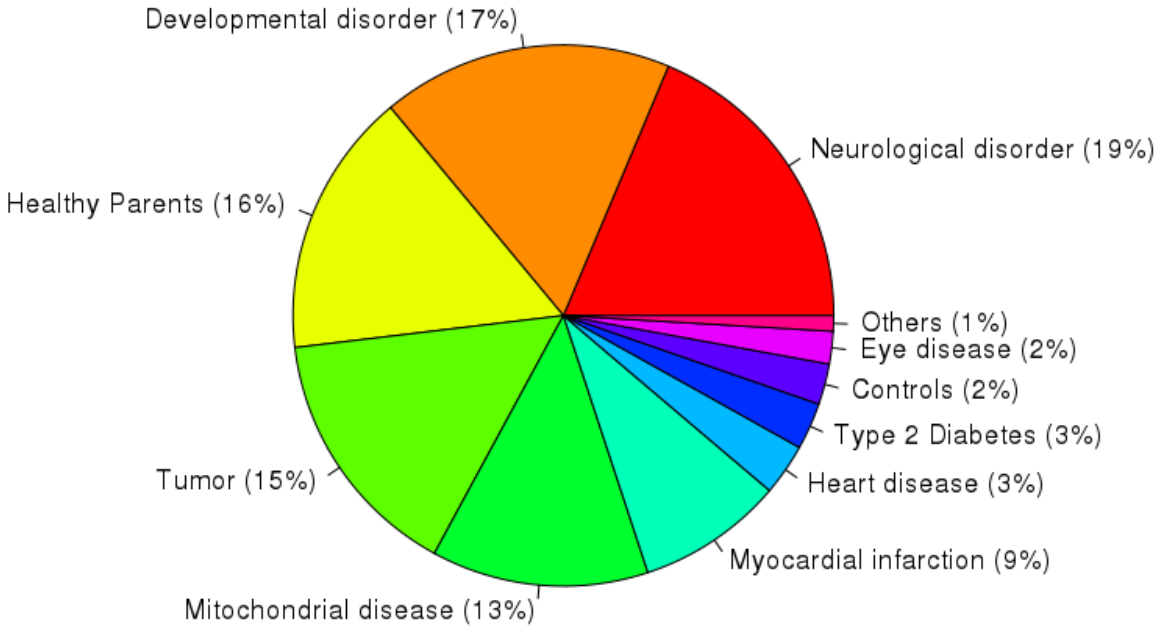
Alice Hadchouel, Thomas Wieland, Matthias Griese, Enrico Baruffini, Bettina Lorenz-Depiereux, Laurent Enaud, Elisabeth Graf, Jean Christophe Dubus, Sonia Halioui-Louhaichi, Aurore Coulomb, Christophe Delacourt, Gertrud Eckstein, Ralf Zarbock, Thomas Schwarzmayr, François Cartault, Thomas Meitinger, Tiziana Lodi, Jacques de Blic, and Tim M. Strom

Figure S1. Haplotype mapping

rsSNP	Chromosome	hg19 position	53654	53655	37601	37602	37603	37604	37605	37606	37607	37609	37610	37611	37612	37613	37614	37615	37617	Genes
rs11171747	chr12	56518408	GG	TT	GG	TT	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	
rs3809134	chr12	56546011	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	
rs12810816	chr12	56552960	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	
rs1290898	chr12	56559840	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	
rs7960225	chr12	56564811	AA	GG	AA	GG	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	
rs1274500	chr12	56658859	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	
rs744051	chr12	56667298	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	
rs10783780	chr12	56704152	AA	AG	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	
rs2066808	chr12	56737973	AA	AG	AA	GG	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	
rs774047	chr12	56815922	TC	TT	TT	TT	TT	TT	TT	TT	TC	TT	TT	TT	TT	TT	TT	TT	TC	
rs774049	chr12	56816978	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	
rs703829	chr12	56823622	TC	TT	TT	TT	TT	TT	TT	TT	TC	TT	TT	TT	TT	TT	TT	TT	TC	
rs774033	chr12	56825311	TC	TT	TT	TT	TT	TT	TT	TT	TC	TT	TT	TT	TT	TT	TT	TT	TC	
rs774039	chr12	56825981	AG	GG	GG	GG	GG	GG	GG	GG	AG	GG	GG	GG	GG	GG	GG	GG	AG	
rs774045	chr12	56838745	AG	GG	AA	GG	AA	AA	AA	AA	AG	AA	AA	AA	AA	AA	GG	GG	AG	
rs1082214	chr12	56846490	TC	CC	TT	CC	TT	TT	TT	TT	TC	TT	TT	TT	TT	TT	CC	CC	TC	
rs941208	chr12	56983252	TT	TC	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	CC	CC	TC	
rs10450	chr12	56984606	TT	TC	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	CC	CC	TC	
rs9368	chr12	56988342	CC	AC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	AA	AA	AC	
rs1465081	chr12	57050174	CC	TC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	TT	TT	TC	
rs2277339	chr12	57146069	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	
rs10876921	chr12	57176186	GG	AG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	AA	AA	AG	
rs10876931	chr12	57212490	GG	AG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	AG	
rs10876933	chr12	57221865	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	AA	AA	GG	
rs1466383	chr12	57231497	CC	TC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	TC	
rs1078043	chr12	57232836	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	
rs898612	chr12	57234406	AA	AG	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AG	
rs4759035	chr12	57255135	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	
rs11832720	chr12	57257244	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	
rs7315229	chr12	57288668	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	
rs1846400	chr12	57293436	AA	AG	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AG	
rs1391708	chr12	57305580	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	
rs10506347	chr12	57331087	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	
rs725957	chr12	57331741	AA	AG	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AG	
rs1072669	chr12	57342188	AA	AG	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	
rs7136770	chr12	57381259	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	
rs733629	chr12	57406444	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	
rs4759272	chr12	57438658	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	
rs1059513	chr12	57489709	TT	TC	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	
rs703817	chr12	57489828	TT	TC	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	CC	CC	TC	
rs324015	chr12	57490100	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	TT	TT	TC	STAT6
rs841718	chr12	57492996	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	GG	GG	AG	STAT6
rs2598483	chr12	57506905	CC	CC	CC	CC	CC	CC	-	CC	CC	CC	CC	CC	CC	CC	CC	CC	TC	
rs324013	chr12	57510661	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	CC	CC	TC	
rs11172113	chr12	57527283	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	LRP1
rs715948	chr12	57532982	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	TT	TT	TC	LRP1
rs10876966	chr12	57543572	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	LRP1
rs1800159	chr12	57593894	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	GG	GG	AG	LRP1
rs10783815	chr12	57616013	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	AA	AA	AG	NXPH4
rs7485577	chr12	57616061	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	
rs7486863	chr12	57681122	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	GG	GG	AA	
rs4760355	chr12	57725197	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	
rs3809114	chr12	57848639	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	AG	
rs2228224	chr12	57865321	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	AG	
rs11544238	chr12	57870155	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AC	ARHGAP9
rs3825080	chr12	57871555	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	ARHGAP9
																				MARS; DDIT3
rs1148556	chr12	57917525	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	MBD6
rs2127318	chr12	57920262	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	GG	GG	TG	
rs1284605	chr12	57921188	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	CC	CC	TC	
rs11172254	chr12	57968738	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	AG	
rs775251	chr12	57978740	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	TC	
rs812315	chr12	57993490	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AG	
rs2277323	chr12	58009372	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	AG	
rs2277324	chr12	58013175	GG	GG	GG	GG	GG	AG	GG	GG	AG	GG	GG	GG	GG	GG	GG	GG	GG	
rs923828	chr12	58015494	GG	GG	GG	GG	GG	AG	GG	GG	AG	GG	GG	GG	GG	GG	GG	GG	GG	
rs715930	chr12	58023981	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	AC	
rs11172300	chr12	58076515	CC	CC	CC	CC	CC	TC	CC	CC	TC	CC	CC	CC	CC	CC	CC	CC	CC	
rs701008	chr12	58117645	CC	CC	CC	CC	CC	TC	CC	CC	TC	CC	CC	CC	CC	CC	CC	CC	CC	
rs4760169	chr12	58118847	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	
rs10877011	chr12	58124992	GG	GG	GG	GG	GG	TG	GG	GG	TG	GG	GG	GG	GG	GG	GG	GG	GG	
rs12368653	chr12	58133256	GG	GG	GG	GG	GG	AG	GG	GG	AG	GG	GG	GG	GG	GG	GG	GG	GG	
rs2069502	chr12	58144665	TT	TT	TT	TT	TT	TC	TT	TT	TC	TT	TT	TT	TT	TT	TT	TT	TT	
rs1048691	chr12	58152948	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	
rs8176345	chr12	58158558	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	
rs703842	chr12	58162739	GG	GG	GG	GG	GG	AG	GG	GG	AG	GG	GG	GG	GG	GG	GG	GG	GG	
rs2291617	chr12	58166403	GG	GG	TG	GG	GG	TG	TG	GG	TG	GG	GG	GG	GG	GG	GG	GG	GG	
rs4760332	chr12	58222672	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	
rs10783853	chr12	58234262	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	
rs7954957	chr12	58299250	TT	TT	TT	TT	TT	TC	TT	TT	TC	TT	TT	TT	TT	TT	TT	TT	TT	

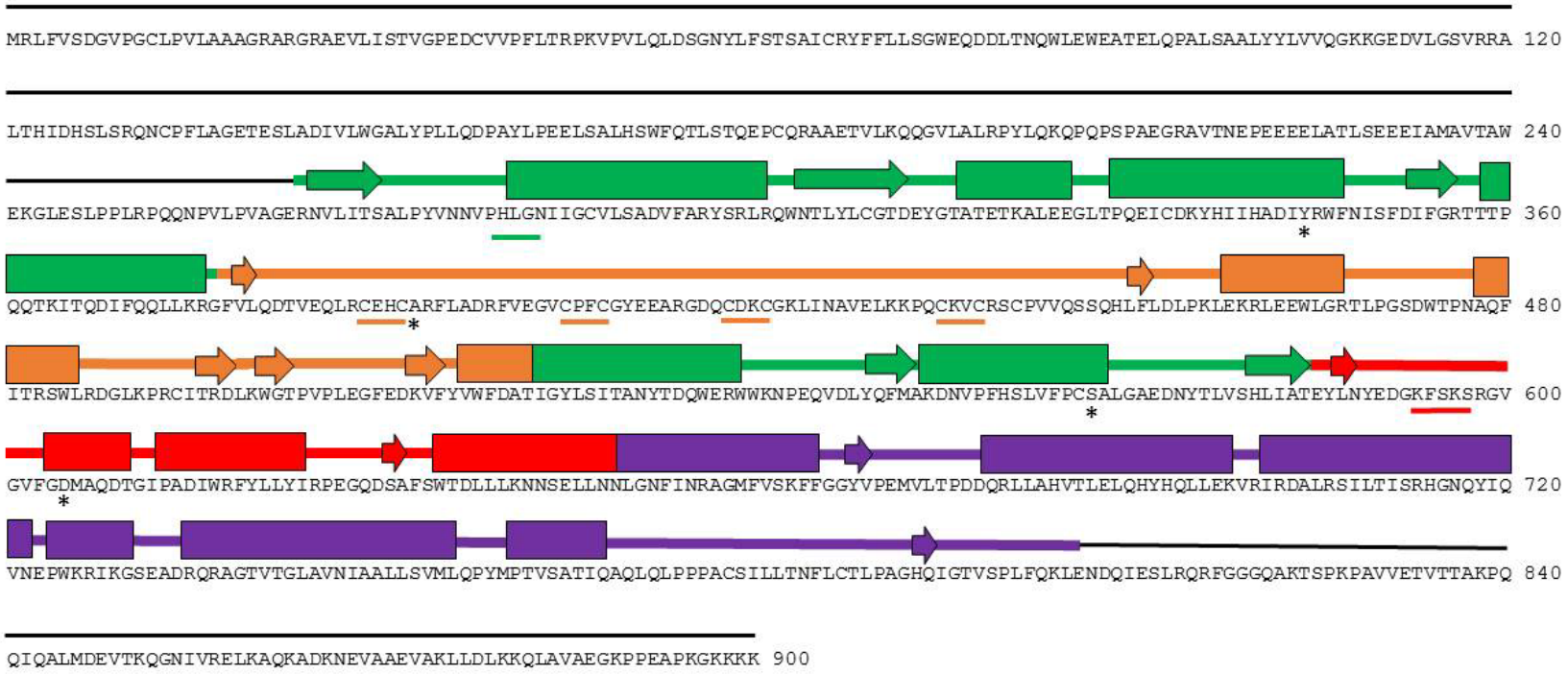
Genotypes common to affected individuals from Réunion are indicated by yellow background. The minimal critical region is located between rs703817 and rs2277324. Genotypes indicated by grey background are of affected individuals from Tunisia and Paris.

Figure S2. In-house exomes



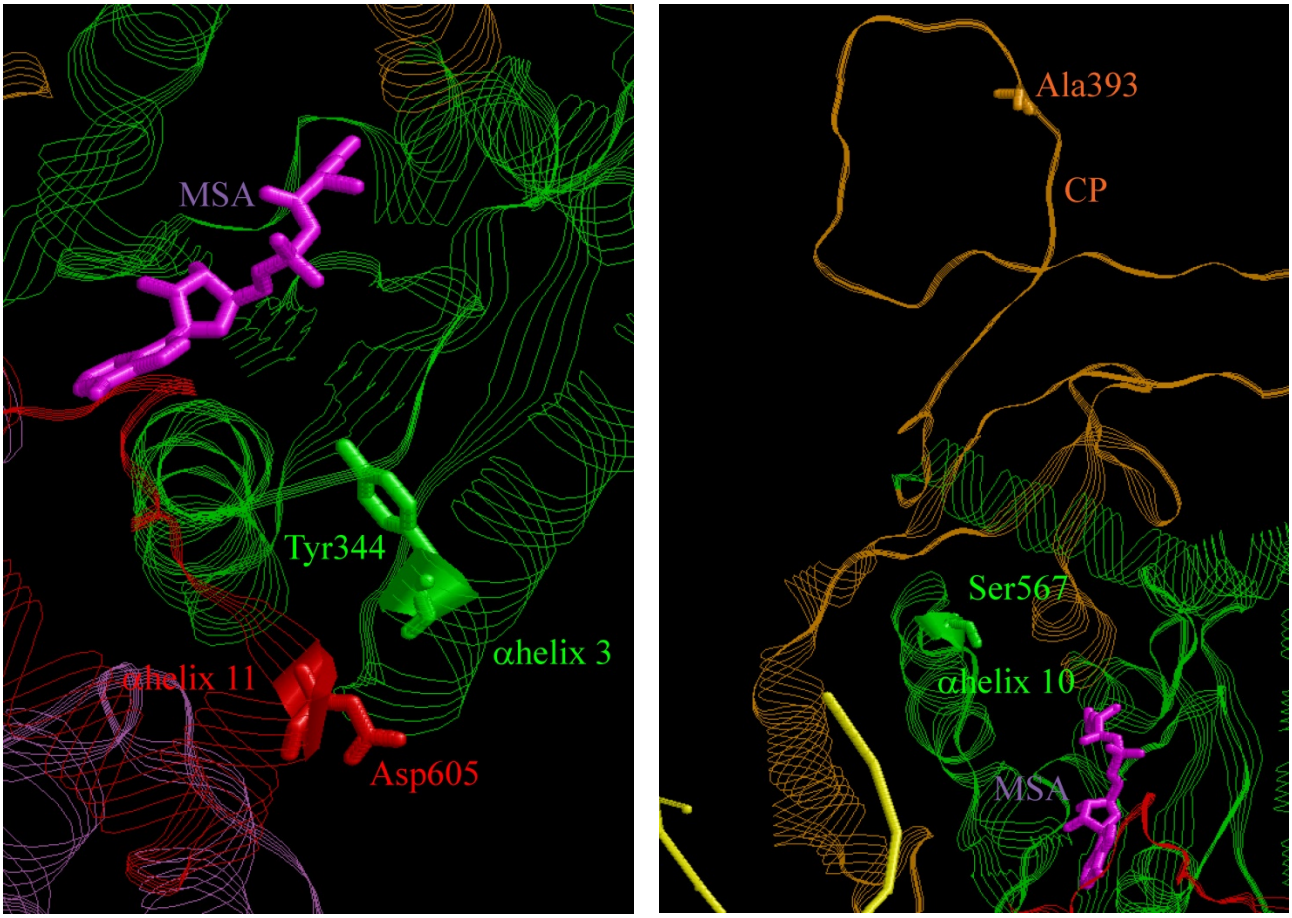
Approximately 4000 in-house exomes of individuals with unrelated rare and complex diseases were used as controls. The pie graph displays the proportion of disease groups.

Figure S3. Secondary structure prediction of human methionyl-tRNA synthetase



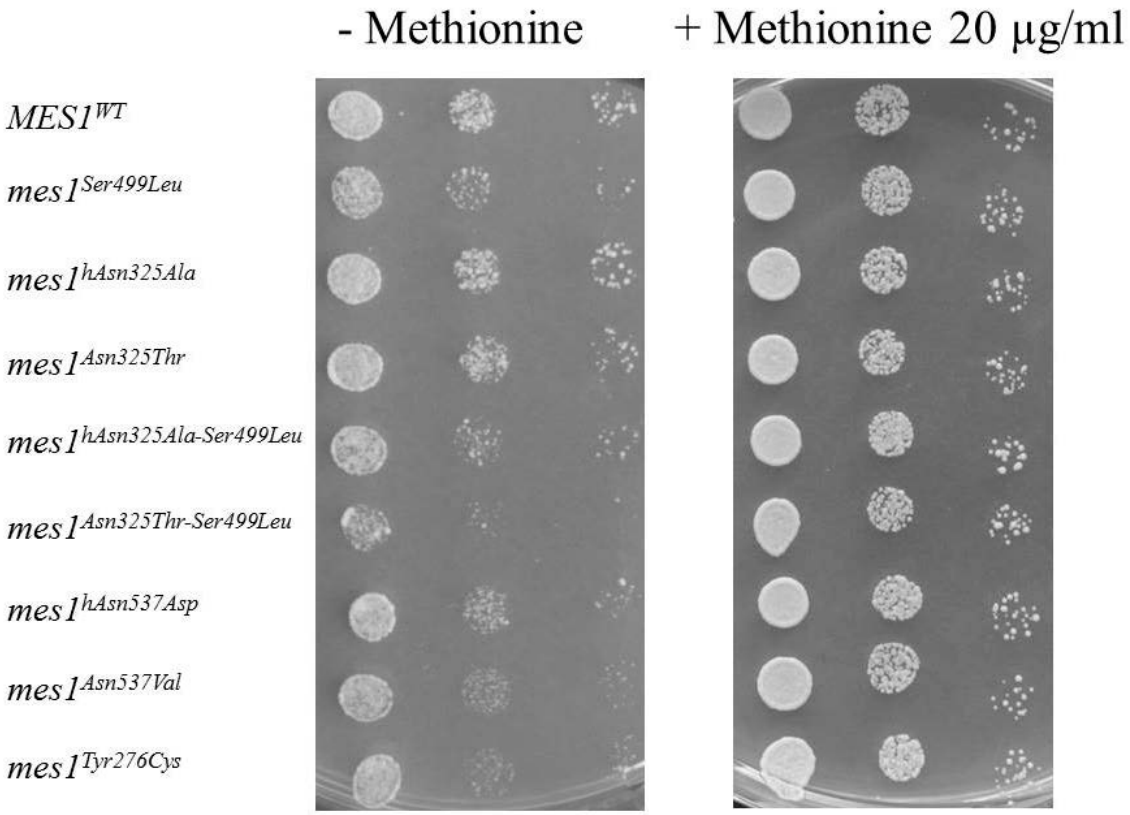
Secondary structural elements are presented above the sequence. Colors: green, the Rossmann fold domain; orange, the CP insertion; red, the stem-contact fold domain; violet, the α-helix bundle domain; black, structure not predicted. Asterisks indicate mutated amino acid positions. The conserved motifs in the catalytic site HIGH and KMSKS are underscored in green and red, respectively. The four CXX[C,D,H] motifs involved in the binding of the zinc ion are underscored in orange[1-3].

Figure S4. Structure of human methionyl-tRNA synthetase around the mutant amino acids



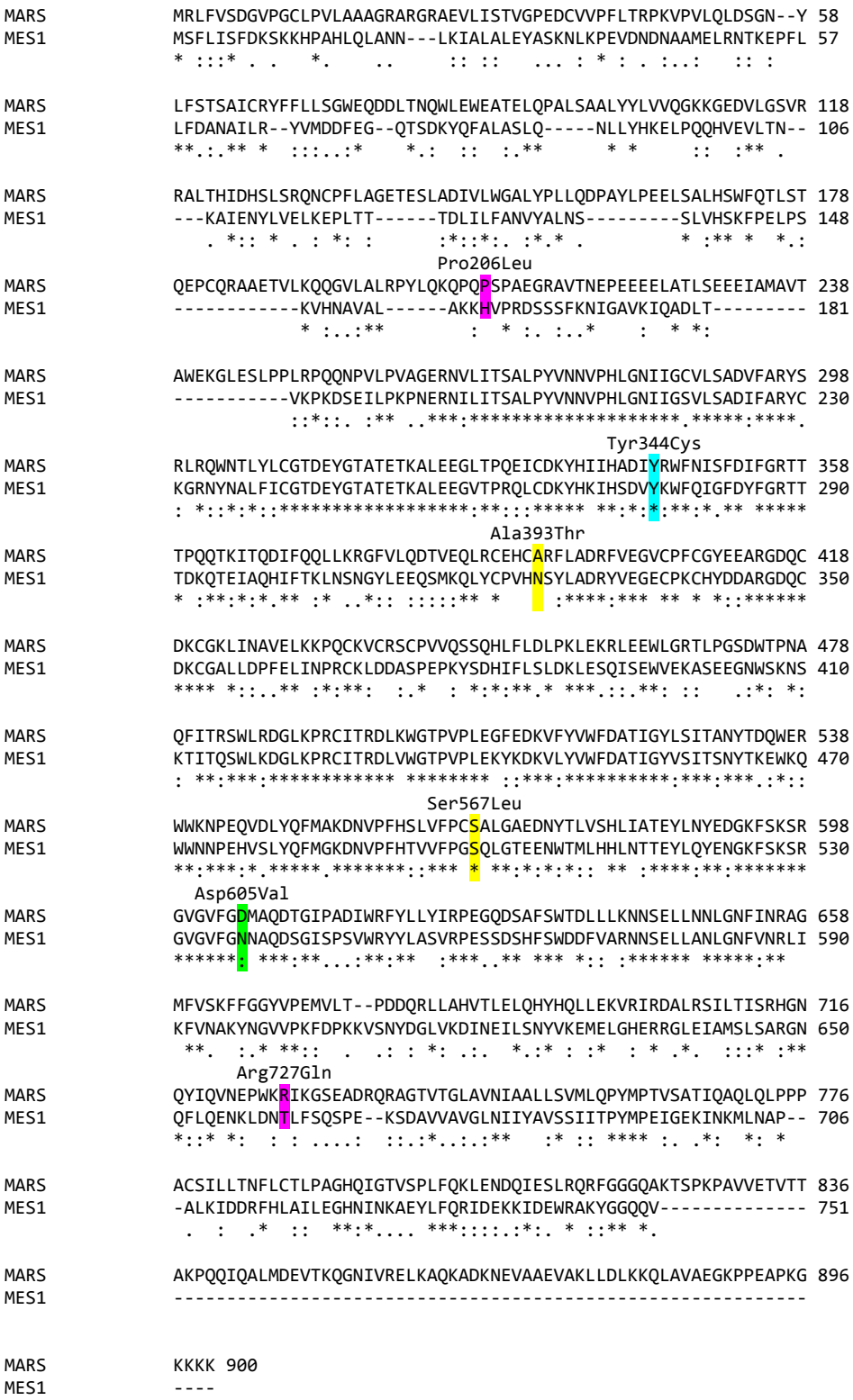
Left: Tyr344, which forms part of the α-helix 3 inside the Rossman fold domain, and Asp605, which forms part of the α-helix 11 inside the stem-contact-fold domain. Right: Ala393, which forms part of a loop inside the highly conserved CP insertion, and Ser567, which forms part of the α-helix 10 inside the Rossman fold domain. MSA: methionyl sulfamoyl adenosine.

Figure S5. **Growth of wt and *mes1* mutant strains**



Growth of wt and *mes1* mutant strains on solid medium without methionine (left panel) or with methionine 20 µg/ml (right panel). Cells were spotted with serial dilutions ($5x10^3$, $5x10^2$, $5x10^1$ cells/spot) and pictures were taken after 36 hours.

Figure S6. Alignment of human MARS and yeast MES1



Sequences NP_004981 and CAA97293.1 were aligned with Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo). * indicates conserved amino acids, : and . indicate conservation between groups with strong and weak similar properties, respectively. The residues Tyr344, Ser567, Asp605 and Ala393 are located in highly conserved regions (86%, 86%, 86% and 50% identity calculated on the basis of the flanking 14 amino acids), whereas the residues Pro206 and Arg727 are located in poorly conserved regions (7% plus gap and 14% identity, respectively).

Table S1. Clinical description of all cases

ID	ID[a]	Sex	Age at onset (m)	Age at diagnosis	Symptoms at diagnosis	Liver disease	Pulmonary fibrosis	WLL	Other treatment	Outcome	Age at death	Age if alive	Outcome if alive
37601	27	m	3	4.2 m	dyspnea, low SaO2, FTT	HMG, high AST, ALT and GGT, steatosis extensive fibrosis	no	yes	no	deceased	17.3 m		
37602		m	2	10 m	dyspnea	NA	undetermined	no	no	deceased	10 m		
37603	19	m	3	7.4 y	dyspnea, cough, digital clubbing	HMG, high AST, ALT and GGT	moderate	yes	iv steroids, hydroxychloroquine cyclophosphamide	deceased	14.3 y		
37604	21	f	9	16.1 m	dyspnea, low SaO2, FTT	HMG, high AST, ALT and GGT,	no	yes	iv steroids	alive		14.9 y	exercise desaturation
37605	28	f	6	22.7 m	dyspnea, cough, FTT	HMG	no	yes	iv steroids	alive		10.2 y	exercise desaturation
37607	20	m	2	11.9 m	dyspnea, low SaO2, digital clubbing, FTT	HMG, high AST, ALT and GGT, cirrhosis	no	yes	no	alive		15.8 y	asymptomatic
37609	24	m	10	2.3 y	dyspnea, low SaO2, digital clubbing	high AST, HMG	no	yes	iv steroids	alive		12 y	CRI, nocturnal oxygen
37610	26	m	6	15.5 m	dyspnea	HMG, high AST, ALT and GGT	diffuse	yes	iv steroids	alive		11.1 y	CRI, continuous oxygen
37611	22	f	3	15.7 m	dyspnea, low SaO2, digital clubbing, FTT	HMG, high AST	no	yes	no	deceased	18.6 m		
37612	25	f	2.5	5.9 m	dyspnea, cough, FTT	high AST, ALT and GGT, HMG	mild	yes	iv steroids	alive		11.5 y	lung transplantation one year ago
37613	15	m	3	3.8 y	dyspnea, low SaO2, digital clubbing, FTT	HMG, high AST	diffuse	yes	iv steroids mycophenolate mofetil	deceased	15.1 y		
53655		m	1	7.8 m	dyspnea, cough, low SaO2, FTT	HMG, high AST, ALT and GGT, cirrhosis	mild	yes	no	alive		4.2 y	asymptomatic
53654	29	m	1	4 m	dyspnea, cough, low SaO2, FTT	HMG, high AST, ALT and GGT, steatosis cirrhosis	diffuse	yes	IV steroids	alive		7.8 y	asymptomatic
69490	11	f	3	22.1 y	dyspnea, digital clubbing	HMG, high GGT	moderate	no	no	alive		24.9 y	CRI, nocturnal oxygen
69502	30	f	1.5	2.9 m	dyspnea, FTT	HMG, high AST and GGT, cirrhosis	mild	yes	iv steroids azathioprine	alive		5.8 y	CRI, nocturnal oxygen

69514	31	m	1.5	3.2 m	dyspnea, FTT	HMG, high AST, ALT and GGT, steatosis cirrhosis	undetermined	yes	no	deceased	3.5 y		
69517	34	m	2	3 m	dyspnea, FTT	HMG, high AST and GGT	undetermined	yes	no	alive		1.1 y	CRI, nocturnal oxygen
69518	32	m	1	1.8 m	dyspnea, FTT	HMG, high AST, ALT and GGT, cirrhosis	moderate	yes	no	alive		4.6 y	asymptomatic
37606	23	m	2	3.7 m	dyspnea, FTT	HMG, high GGT, steatosis cirrhosis	undetermined	yes	no	deceased	15.5 m		
P1	10	f	72	16.3 y	dyspnea, digital clubbing	high GGT, HMG	diffuse	no	oral steroids	deceased	25.2 y		
P2	18	m	3	3.5 m	dyspnea, low SaO2, FTT	HMG, high GGT steatosis	undetermined	yes	no	deceased	6.9 m		
P3	33	m	1.5	4.4 m	dyspnea, cough, low SaO2, FTT	HMG, high AST and GGT	undetermined	yes	iv steroids	deceased	5.4 m		
P4		f	2	9.4 m	dyspnea, low SaO2, FTT	HMG, high GGT, cirrhosis	diffuse	yes	iv steroids	alive		4.7 y	asymptomatic
P5		m	2	4.4 m	dyspnea, low SaO2, FTT	HMG, high AST, ALT	undetermined	yes	no	deceased	8.5 m		
P6		m	2.5	5.5 m	dyspnea, FTT	HMG, high AST, ALT and GGT	undetermined	yes	no	alive		5.2 y	No respiratory symptoms malnutrition
P7		m	0.5	3.6 m	dyspnea, cough, FTT	HMG, high AST, ALT and GGT	undetermined	yes	iv steroids	alive		1.3 y	CRI, continuous oxygen
37614		m		3.6 y	Exercise desaturation	None	undetermined	yes	no	alive		18.1 y	Asymptomatic
37615		f	46	4.9 y	dyspnea, low SaO2, digital clubbing, FTT	None	undetermined	yes	no	deceased	9.9 y		
37617		m	10	12 m	dyspnea, low SaO2, FTT	None	diffuse	yes	iv steroids, cyclophosphamide	alive		22.3 y	Exercise dyspnea

Abbreviations: m: male; f: female; SaO_2: oxygen saturation; FTT: failure to thrive; CRI: chronic respiratory insufficiency; HMG: hepatomegaly; m: months; y: years; WLL: whole lung lavages; iv: intra-venous; NA: not available; [a]ID used in Enaud et al[4]. 69514 and 69517, 69518 and 37606, 37614 and 37615 are sibs.

The initial clinical course of the disease in the Tunisian sib pair (37614 and 37615) has been described elsewhere[5]. The 4.1 year old girl presented with chronic cough, dyspnea and nail clubbing. Chest X-ray and CT showed a reticulonodular pattern. Lung lavage was milky and cytology was characteristic for PAP. During episodes of respiratory distress, she was treated with whole lung lavages. She died at the age of 9.9 years because of respiratory insufficiency. Her brother had dyspnea and desaturation during exercise at the age of 3 years. Diagnosis of PAP was suspected because of the disease of his sister and proven by open lung biopsy. Now, at the age of 18 years, he is free of respiratory distress.

The boy from metropolitan France (37617) was diagnosed with PAP by open lung biopsy at the age of 12 months. Several lung lavages were performed from diagnosis till the age of 8 years. By the age of 9 years, the disease advanced to lung fibrosis and he was treated multiple times

with high-dose IV steroids and cyclophosphamide. He was treated with oral steroids from the age 13 to 18 years. He is now 23 years old and suffers from exercise dyspnea. He has no specific treatment and does not need supplemental oxygen. The last CT showed bilateral, diffuse reticulonodular syndrome with septal thickening and large cystic areas. His last lung function test showed a total lung capacity (TLC) of 73%, forced vital capacity (FVC) of 64%, FEV1/FVC ratio of 94%, diffusing capacity (DLCO) of 66%, but no exercise desaturation.

Table S2. Analyses performed in 29 individuals with PAP and genotypes identified in *MARS*

ID	Origin	c.1177G>A p.Ala393Thr	c.1700C>T p.Ser567Leu	c.1814A>T p.Asp605Val	c.1031A>G p.Tyr344Cys	Mother heterozygous	Father heterozygous	Linkage	Exome	Genome	Sanger
37601	Reunion	A/A	T/T			x	x	x			x
37602	Reunion	A/A	T/T	A/A	A/A			x	x		x
37603	Reunion	A/A	T/T					x			x
37604	Reunion	A/A	T/T			x	x	x			x
37605	Reunion	A/A	T/T			x	x	x			x
37607	Reunion	A/A	T/T	A/A	A/A	x	x	x	x	x	x
37609	Reunion	A/A	T/T					x			x
37610	Reunion	A/A	T/T			x	x	x			x
37611	Reunion	A/A	T/T					x			x
37612	Reunion	A/A	T/T			x	x	x			x
37613	Reunion	A/A	T/T					x			x
53655	Comoros	A/A	T/T					x			x
53654	Reunion	A/A	T/T					x			x
69490	Reunion	A/A	T/T								x
69502	Reunion	A/A	T/T			x	x				x
69514	Reunion	A/A	T/T	A/A	A/A	x	x			x	x
69517	Reunion	A/A	T/T			x	x				x
69518	Reunion	A/A	T/T			x	x				x
37606	Reunion	A/A	T/T			x	x	x			x
P1	Reunion	A/A	T/T	A/A	A/A				x		x
P2	Reunion	A/A	T/T	A/A	A/A				x		x
P3	Reunion	A/A	T/T	A/A	A/A				x		x
P4	Comoros	A/A	T/T								x
P5	Comoros	A/A	T/T								x
P6	Comoros/Madagascar	A/A	T/T								x
P7	Reunion	A/A	T/T	A/A	A/A				x		x
37614	Tunisia	G/G	C/C	T/T		x	x	x			x
37615	Tunisia	G/G	C/C	T/T	A/A	x	x	x	x		x
37617	Reunion/France	G/A	C/T	A/A	A/G	x	x	x	x		x

Table S3. Non-synonymous variants in the critical region between markers rs703817 and rs2277324 in 2 affected individuals from Reunion Island

hg19 position	ref allele	alt allele	Gene	Function	In-house controls hom ref—het—hom alt	ExAC – European origin hom ref—het—hom alt	ExAC – African origin hom ref—het—hom alt
chr12 57589784	A	C	LRP1	missense	8-6-5319	0-5-34294	12-333-4948
chr12 57619362	G	A	NXPH4	missense	2419-2342-572	14353-15207-4037	2369-2264-518
chr12 57894189	G	A	MARS	missense	5333-0-0	34426-0-0	5331-4-0
chr12 57906083	C	T	MARS	missense	5333-0-0	34424-0-0	5331-4-0
chr12 57979190	C	G	BC033961	missense	3399-1588-346	-	-

Table S4. Oligonucleotides used for the yeast studies

Oligonucleotide	Sequence
MES1CBamFw	GGCCGGGAtcccctcactaactttgtgg[a]
MES1CPstIRv	CGGGGCTGCAGctcatttcaagcgacgcagg[a]
MES1DFw	tctttcctcatttcctttgataaatcgaagaaacatcctgccGCTTCGTACGCTGCAGGTCGACG[b]
MES1DRv	agaggaaaaaaaatgcttccgatgaaccttacacttgttgaccGATATCATCGATGAATTCGAGC[b]
MESPreXbaFw	ggtactgatgaatatggtactgc
MESPostXbaRv	ctggggactgtgaaaacaagg
MES1hN325AFw	gtactgtccagttcatGCttcttatctggctgatcgttacgtgg[c]
MES1hN325ARv	ccacgtaacgatcagccagataagaaGCatgaactggacagtac[c]
MES1N325TFw	gtactgtccagttcataCttcttatctggctgatcgttacgtgg[c]
MES1N325TRv	ccacgtaacgatcagccagataagaaGtatgaactggacagtac[c]
MES1S499LFw	ccatacagttgttttccctggtCTAcaattgggtacggaagag[c]
MES1S499LRv	ctcttccgtacccaattgTAGaccagggaaaacaactgtatgg[c]
MES1hN537DFw	gtaggggtgttggtgtttttggtGataacgctcaagactctgg[c]
MES1hN537DRv	ccagagtcttgagcgttatCaccaaaaacaccaacacccctac[c]
MES1N537VFw	gtaggggtgttggtgtttttggtGTtaacgctcaagactctgg[c]
MES1N537VRv	ccagagtcttgagcgttaACaccaaaaacaccaacacccctac[c]
MES1Y276CFw	cacaaaatccacagtgacgtttGcaagtggttccaaattggatttg[c]
MES1Y276CRv	caaatccaatttggaaccacttgCaaacgtcactgtggattttgtg[c]

[a]In uppercase the clamp and the restriction site, in lowercase the sequence flanking the *MES1* gene.
[b]In lowercase the sequence flanking the *MES1* open reading frame, in uppercase the sequences flanking the KanMX4 cassette.
[c]In upper case the mutant nucleotides.

References

1. Mechulam, Y., Schmitt, E., Maveyraud, L., Zelwer, C., Nureki, O., Yokoyama, S., Konno, M., and Blanquet, S. (1999). Crystal structure of Escherichia coli methionyl-tRNA synthetase highlights species-specific features. J Mol Biol *294*, 1287-1297.
2. Schmitt, E., Meinnel, T., Blanquet, S., and Mechulam, Y. (1994). Methionyl-tRNA synthetase needs an intact and mobile 332KMSKS336 motif in catalysis of methionyl adenylate formation. J Mol Biol *242*, 566-576.
3. Casina, V.C., Lobashevsky, A.A., McKinney, W.E., Brown, C.L., and Alexander, R.W. (2011). Role for a conserved structural motif in assembly of a class I aminoacyl-tRNA synthetase active site. Biochemistry *50*, 763-769.
4. Enaud, L., Hadchouel, A., Coulomb, A., Berteloot, L., Lacaille, F., Boccon-Gibod, L., Boulay, V., Darcel, F., Griese, M., Linard, M., et al. (2014). Pulmonary alveolar proteinosis in children on La Reunion Island: a new inherited disorder? Orphanet J Rare Dis *9*, 85.
5. Halioui-Louhaichi, S., Ben Hariz, M., Hamzaoui, A., Mestiri, T., Jerbi, E., Louzir, B., Ben Khelifa, S., Ben Jaballah, N., De Blic, J., Ben Ammar, M.S., et al. (2005). Pulmonary alveolar proteinosis in two siblings. A case report. Tunis Med *83*, 488-491.