

Title

Clinical, biochemical, and genetic heterogeneity in glutaric aciduria type II patients

Amanat Ali ¹, Fatmah Saeed Ali Almesmari ¹, Nahid Al Dhahouri ¹, Arwa Mohammad Saleh Ali ¹, Mohammed Ahmed Ali Mohamed Ahmed Aldhanhani ¹, Ranjit Vijayan ², Amal Al Tenaiji ³, Aisha Al Shamsi ⁴, Jozef Hertecant ⁴ and Fatma Al Jasmi ^{1,4*}

¹Department of Genetics and Genomics, College of Medicine and Health Sciences, United Arab Emirates University, PO Box 15551, Al Ain, Abu Dhabi, United Arab Emirates

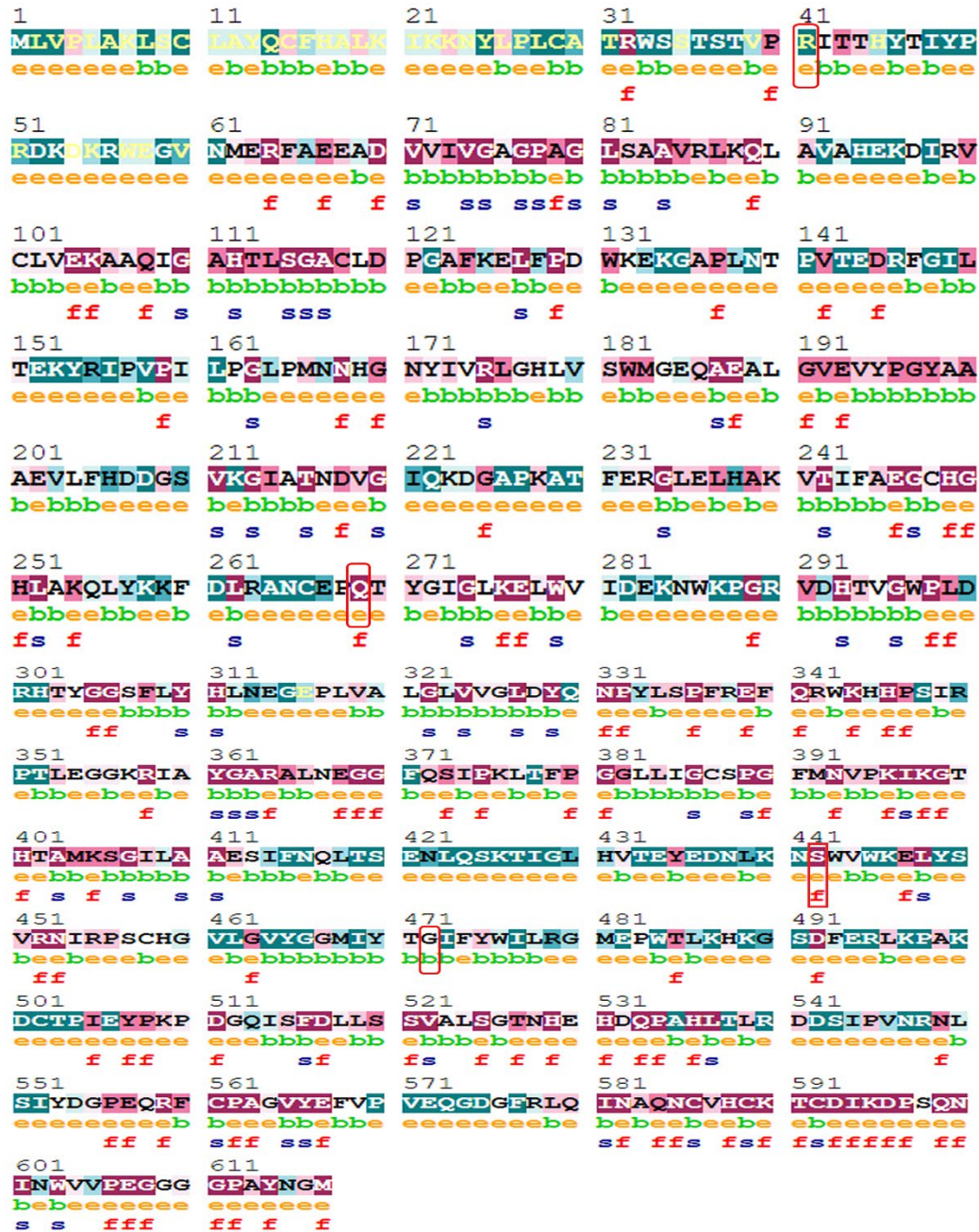
²Department of Biology, College of Science, United Arab Emirates University, PO Box 15551, Al Ain, Abu Dhabi, United Arab Emirates

³Department of Pediatrics, Sheikh Khalifa Medical City, PO Box 51900, Abu Dhabi, United Arab Emirates

⁴Department of Pediatrics, Tawam Hospital, PO Box 15551 Al Ain, United Arab Emirates

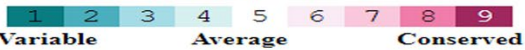
*Correspondence: aljasmi@uaeu.ac.ae; Tel.: +971-3-7137412; Fax: +971-3-7672022

Supplementary Information



Legend:

The conservation scale:



e - An exposed residue according to the neural-network algorithm.

b - A buried residue according to the neural-network algorithm.

f - A predicted functional residue (highly conserved and exposed).

s - A predicted structural residue (highly conserved and buried).

x - Insufficient data - the calculation for this site was performed on less than 10% of the sequences.

Figure S1. Conservation of amino acid predicted by ConSurf. Square box shown in red color represents the location of mutation.

Table S1. MutPred2 based prediction of molecular mechanisms affected by missense variants.

Variant	Molecular mechanism	MutPred2 score/g scores	P-value	Interpretation
<i>ETFDH</i> :p.Arg41Leu	Loss of glycosylation at S37	0.566	0.02	Disease associated, actionable hypothesis
	Loss of disorder	0.566	0.03	Disease associated, actionable hypothesis
<i>ETFDH</i> :p.Gln269His	Altered Ordered interface	0.869	0.02	Disease associated, confident hypothesis
	Loss of Allosteric site at Y271	0.869	7.4e-03	Disease associated, very confident hypothesis
	Altered Metal binding	0.869	0.04	Disease associated, confident hypothesis
	Loss of Catalytic site at Y271	0.869	0.01	Disease associated, confident hypothesis
	Altered DNA binding	0.869	0.01	Disease associated, confident hypothesis
	Altered Transmembrane protein	0.869	0.03	Disease associated, confident hypothesis
	Loss of Relative solvent accessibility	0.869	0.03	Disease associated, confident hypothesis
<i>ETFDH</i> :p.Ser442Leu	Loss of ubiquitination at K440	0.881	0.04	Disease associated, confident hypothesis
	Loss of disorder	0.881	0.07	Disease associated
<i>ETFDH</i> :p.Gly472Arg	Gain of molecular recognition features (MoRFs) binding	0.985	0.004	Disease associated, very confident hypothesis
	Gain of methylation at G472	0.985	0.03	Disease associated, confident hypothesis

Threshold P value ≤ 0.05 , Interpretation is performed on the basis of g and P scores; actionable hypotheses: $g > 0.5$, $P < 0.05$; confident hypotheses: $g > 0.75$, $P < 0.05$ and very confident hypotheses: $g > 0.75$, $P < 0.01$. Disease associated when $g > 0$.

Table S2. Effect of variants on the structure and domains of ETF-QO predicted by HOPE.

Variant	Amino acid properties	Effect on protein structure and domain
<i>ETFDH</i> :p.Arg41Leu	<p>There is a difference in charge between the wild-type and mutant amino acid. The charge of the wild-type residue will be lost, this can cause loss of interactions with other molecules or residues.</p> <p>The wild-type and mutant amino acids differ in size. The mutant residue is smaller, this might lead to loss of interactions.</p> <p>The mutation introduces a more hydrophobic residue at this position. This can result in loss of hydrogen bonds and/or disturb correct folding.</p>	<p>The mutated residue is located in a domain that is important for the main activity of the protein.</p> <p>Mutation of the residue might disturb this function.</p>
<i>ETFDH</i> :p.Gln269His	<p>The wild-type and mutant amino acids differ in size.</p> <p>The mutant residue is bigger, this might lead to bumps.</p>	<p>The mutated residue is located in a domain that is important for the main activity of the protein.</p> <p>Mutation of the residue might disturb this function.</p>
<i>ETFDH</i> :p.Ser442Leu	<p>The wild-type and mutant amino acids differ in size.</p> <p>The mutant residue is bigger, this might lead to bumps.</p> <p>The mutation introduces a more hydrophobic residue at this position. This can result in loss of hydrogen bonds and/or disturb correct folding.</p>	<p>The residue is located in a domain that was annotated as an intramembrane region.</p> <p>The wildtype residue was probably making hydrophilic interactions with other residues. The mutation could disturb these interactions and affect the structure of the intramembrane domain.</p> <p>The mutant residue Leu442 will probably not fit here and bump into other residues and/or lipid molecules, thereby could disturb the structure of the intramembrane domain.</p> <p>The mutated residue is located in a domain that is important for the main activity of the protein.</p> <p>Mutation of the residue might disturb this function.</p>
<i>ETFDH</i> :p.Gly472Arg	<p>There is a difference in charge between the wild-type and mutant amino acid.</p> <p>The mutation introduces a charge, this can cause repulsion of ligands or other residues with the same charge. The wild-type and mutant amino acids differ in size.</p> <p>The mutant residue is bigger, this might lead to bumps.</p> <p>The torsion angles for this residue are unusual. Only glycine is flexible enough to make these torsion angles, mutation into another residue will force the local backbone into an incorrect conformation and will disturb the local structure.</p>	<p>The wild-type residue is a glycine, the most flexible of all residues. This flexibility might be necessary for the protein's function.</p> <p>Mutation of this glycine can abolish this function</p> <p>The mutated residue is located in a domain that is important for the main activity of the protein.</p> <p>Mutation of the residue might disturb this function.</p>

Table S3. UniProt Accession number of sequences used for protein sequence alignment.

Name of organism	Protein ETF-QO
<i>Homo sapiens</i> (Human)	Q16134
<i>Pan troglodytes</i> (Chimpanzee)	H2RB83
<i>Mesocricetus auratus</i> (Hamster)	A0A1U7QWG6
<i>Mus musculus</i> (Mouse)	Q921G7
<i>Rattus norvegicus</i> (Rat)	Q6UPE1
<i>Oryctolagus cuniculus</i> (Rabbit)	U3KM78
<i>Felis catus</i> (Cat)	U5Q9M7
<i>Vulpes vulpes</i> (Fox)	A0A3Q7SNG5
<i>Equus caballus</i> (Horse)	A0A3Q2HQA1
<i>Sus scrofa</i> (Pig)	P55931
<i>Myotis lucifugus</i> (Bat)	G1PDD7
<i>Capra hircus</i> (Goat)	A0A452FRH5
<i>Bos taurus</i> (Bovine)	Q2KIG0
<i>Loxodonta africana</i> (African Elephant)	G3TD93
<i>Ailuropoda melanoleuca</i> (Panda)	G1L6D0
<i>Gallus gallus</i> (Chicken)	F1NY29
<i>Alligator mississippiensis</i> (American alligator)	A0A151PAY4
<i>Physeter macrocephalus</i> (Whale)	A0A2Y9F9E2
<i>Callorhinchus milii</i> (Shark)	V9KD07
<i>Drosophila melanogaster</i> (Fruit fly)	Q7JWF1