

ORIGINAL ARTICLE

Quaternary protein modeling to predict the function of DNA variation found in human mitochondrial cytochrome *c* oxidase

Martin Patrick Horan¹, Jon N Rumbley², Richard G Melvin³, David G Le Couteur⁴ and J William O Ballard¹

Cytochrome *c* oxidase (COX) of the electron transport system is thought to be the rate-limiting step in cellular respiration and is found mutated in numerous human pathologies. Here, we employ quaternary three-dimensional (3-D) modeling to construct a model for human COX. The model was used to predict the functional consequences of amino-acid mutations based on phylogenetic conservation of amino acids together with volume and/or steric perturbations, participation in subunit–subunit interfaces and non-covalent energy loss or incompatibilities. These metrics were combined and interpreted for potential functional impact. A notable strength of the 3-D model is that it can interpret and predict the structural consequences of amino-acid variation in all 13 protein subunits. Importantly, the influence of compensatory changes can also be modeled. We examine mutations listed in the human mutation database Mitomap, and in 100 older men, and compare the results from the 3-D model against the automated MutPred web application tool. In combination, these comparisons suggest that the 3-D model predicts more functionally significant mutations than does MutPred. We conclude that the model has useful functional prediction capability but may need modification as functional data on specific mutations becomes known.

Journal of Human Genetics (2013) 58, 127–134; doi:10.1038/jhg.2012.144; published online 10 January 2013

Keywords: COX; functional prediction; mitochondria; modeling; mtDNA mutation

INTRODUCTION

Cytochrome *c* oxidase (COX, complex IV) of the mitochondrial electron transport system catalyzes the reduction of molecular oxygen¹ and is hypothesized to be the rate-limiting step in oxidative phosphorylation.² COX consists of two copper binding sites (Cu_A and Cu_B), two hemes (a and a₃), as well as a zinc and a magnesium ion. With the transfer of each electron, two protons are pumped through COX from the inner mitochondrial membrane to the intermembrane space. In mammals, COX is composed of 13 subunits consisting of 3 mitochondrial DNA (mtDNA) and 10 nuclear-encoded genes. The three genes expressed from the mtDNA (*MT-CO1*, *MT-CO2* and *MT-CO3*) are the largest of the subunits and form the catalytic core of the enzyme.³ The 10 nuclear-encoded COX genes form the remainder of the enzyme and are directly regulated by nuclear respiratory factors 1 and 2.^{4,5} In mammals, Osada and Akashi⁶ present evidence for ‘compensatory’ adaptive substitutions in nuclear DNA-encoded mitochondrial proteins to prevent fitness decline in primate COX.⁶ However, the frequency of such compensatory mutations within humans is currently unknown at least partly because a robust model of human COX has not been developed. A goal of this study is to develop a robust three-

dimensional (3-D) model for human mitochondrial COX. We then map known mutations onto the model and determine the frequency of mutations in an aged population that are predicted to be functionally significant.

Mutations in the mtDNA-encoded COX genes have been experimentally associated with COX deficiency in various clinical phenotypes, including motor neuron disease,⁷ mitochondrial encephalomyopathy,⁸ mitochondrial myopathy,^{9,10} myoglobinuria,¹¹ mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like (MELAS) episodes,¹² and have been associated with aging¹³ and prostate cancer.¹⁴ Mutations in mtDNA-encoded COX genes can affect single tissues or multiple systems with the severity of COX deficiency varying in different tissues.¹⁵ Further, the severity of COX deficiency may be associated with heteroplasmic load (the presence of more than one mtDNA type) with mitochondrial dysfunction occurring when one mtDNA type harboring a deleterious mutation reaches a threshold level.^{16–18}

In concert with experimental studies, bioinformatic tools have been developed to predict the significance of amino-acid variation on protein structure and function.^{19–21} Many of these algorithms use phylogenetic information to base their predictions on the effect of

¹School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, NSW, Australia; ²Pharmacy Practice and Pharmaceutical Sciences, University of Minnesota-Duluth, Duluth, MN, USA; ³Institute of Biotechnology, University of Helsinki, Helsinki, Finland and ⁴ANZAC Medical Research Institute, Centre for Education and Research on Aging, University of Sydney, Concord RG Hospital, Sydney, NSW, Australia
Correspondence: Dr MP Horan, School of Biotechnology and Biomolecular Sciences, University of New South Wales, High Street, Sydney, NSW 2052, Australia.
E-mail: m.horan@unsw.edu.au

Received 5 August 2012; revised 6 November 2012; accepted 20 November 2012; published online 10 January 2013

amino-acid variation within the context of multiple sequence alignment. 3-D modeling is a further development that can be used to determine the specific interaction of a mutant residue with neighboring residues within individual proteins and protein complexes.^{20,22} Thus, bioinformatic protein modeling is a powerful strategy that can be utilized to make informed predictions into the functional importance of protein amino-acid variation. In mammals, the X-ray crystallography-solved structure for COX has been derived from bovine heart tissue²³ and can be used to form the foundation for generating a model of the complete 13 subunit human COX structure.

In a previous study, we used the homology modeling approach to develop a 3-D model of COX for the fly *Drosophila simulans*.²² This approach enabled us to investigate the structural and functional consequences of a naturally occurring deletion in a nuclear-encoded subunit. Here we use a similar approach to develop an improved quaternary 3-D model for human mitochondrial COX using the Modeller 9.10 program and protein template PDB: 3ASO.^{24,25} We then compared the 3-D model against the MutPred web application tool.²⁶ MutPred was developed to classify an amino-acid substitution as disease-associated or neutral in humans. First, we modeled COX variants listed in the human mutation database, Mitomap.²⁷ This focused approach allowed us to model and make predictions on known functional COX mutations and to quantify the likelihood of variants altering structural conformation and potentially being functionally important. Next, we sequenced the three mtDNA COX subunits from 100 individuals randomly selected from an aged cohort. In combination, we suggest that the model has the potential to be useful for predicting the functional consequences of COX mutations; however, the model may need modification as functional data on specific mutations becomes known. This study did not include an analysis of any mutations found in the COX assembly genes.

MATERIALS AND METHODS

Quaternary 3-D model

The 3-D model of human COX was generated by homology modeling to bovine COX (PDB: 3ASO, 2.30 Å resolution) using Modeller 9.10.^{24,25} The human and bovine (crystallographically resolved) sequence alignment was extracted and the unmatched termini of the human sequence trimmed to match the bovine termini. For example, the first 25 amino acids from COX4I1, the first 45 amino acids from COX5A and the first 31 amino acids from COX6A1 were all trimmed before analysis. This alignment was used to generate three monomeric multisubunit human COX models, including hemes a and a₃, copper centers and additional metals resolved in the bovine structure. The high-resolution lipid structures present in bovine COX were not modeled.

The only structural metric that was not optimized was a small fraction of non-standard side-chain rotomers, an outcome of minimization not deemed critical for the qualitative mutant evaluations here.

Structural parameters measured for functional prediction

The qualitative functional impact of each identified mutation was predicted from four unique parameters. The structure-based parameters included sequence conservation derived from the multiple sequence alignment of 16 mammal and four additional vertebrate sequences. This was combined with the three parameters of local volume or steric perturbation, subunit–subunit interaction interface perturbation and non-covalent energy of interaction loss or incompatibility.

Sequence conservation was determined from the multiple sequence alignment of all COX sequences from 20 species (Supplementary Table 1). More distantly related sequences were avoided to reduce the potential for misinterpretation of conservation arising from second site revertants. Sequences were aligned using PROMALS for each of the 13 COX subunits with the inclusion of bovine complete and crystallographically resolved sequences.²⁸ For

each mutant position analyzed, the position-specific percent identity was determined across the 20 sequences. Because of the limited number of sequences in the alignments and the high conservation of COX, a stringent criteria for mutant impact was used. A mutant was interpreted as having high potential impact if the respective sequence position was 95–100% conserved (exact residue), while a mild impact was assigned if the conservation was between 75 and 94%. Residues providing <74% conservation were considered to be of little impact.

The internal structural parameters were based on a 5 Å interaction sphere around the site of mutation. This volume represents a conservative distance for all non-covalent and nearby covalent residue contacts. The distance was based on an empirically determined van der Waals contact distance for methane of 4.01 Å.²⁹ Local volume or steric effects were evaluated based on the apparent density of side-chain and/or backbone packing within this sphere, indicated by the number of residues within the interaction distance. The potential impact of additional steric restraints imposed by the backbone trajectory was also interpreted (for example, proline angle restriction). A significant impact was assigned when a side-chain methyl group was removed from a densely packed environment. This type of change would lead to cavity formation and side-chain repacking, which can be energetically costly.^{30,31}

Subunit–subunit contacts were also determined within the 5 Å radius about each mutation. A mild impact in this case was assigned if only one or a few atoms from the subunit interface were in contact. Typically, contact across subunit interfaces constituted significantly greater numbers of atoms and residues and the impact of a change was considered greater.

Non-covalent energy perturbations were evaluated based on the mutant side-chain environment and its characteristics (for example, hydrophobic, polar, electrostatic). Included in this category were mutant residues having an immediate impact/perturbation on the COX hemes or other metal centers. A significant impact was assigned when an unpaired charge was generated in an otherwise hydrophobic core.

Ranking the functional significance of the mutations

For each parameter, the relative impact of mutation was evaluated and scored as having no impact (0), mild impact (0.5) or substantial impact (1) across the four parameters. The cumulative score was interpreted as the likelihood of any given mutation to affect the function of COX; that is, '0 + 0 + 0.5 + 1' (total score of 1.5) is very unlikely to affect function, while '1 + 1 + 1 + 1' (a maximum score of 4) is very likely to lead to a functional defect. A predictive ratio score between 0 and 1 was then derived from the total cumulative score and the maximum score to determine functionality. In this study, mutations generating normalized ratio scores of 0.6 and above were considered functionally important. However, this score may need to be modified as additional functional data on the significance of specific mutations becomes known.

COX variants listed in the human mutation database

First, we modeled COX variants listed in the human mutation database Mitomap.²⁷ It is important to note that the variants listed in Mitomap do not represent a random sample of the human population. Further, a substantial number of variants are not reported in the literature and represent direct submissions without supporting correlative data.

We then tested whether mutations predicted by the computer program MutPred have any significant effects on function.²⁶ Mutations generating MutPred scores above 0.5 and assigned with AH (actionable hypothesis, may have influence on disease), CH (confident hypothesis, confidence that variant is disease causing) or VCH (very confident hypothesis, high confidence level that variant is disease causing) were considered to be functional.²⁶

Assaying COX variation from an aged cohort

To determine the frequency of mutations predicted to be functional by the 3-D model and by MutPred, we isolated DNA and sequenced the three core COX genes from 100 aged males. Males were randomly selected from the Concord Health and Aging in Men Project (CHAMP) based in Sydney, Australia.³² Aged individuals were chosen because any identifiable COX variant in this study population is unlikely to be highly deleterious given that these individuals are

long-lived. Overall, the cohort that we studied was age 77 ± 5.0 years, weight 79.8 ± 12.3 kg and body mass index 28.1 ± 4.1 . DNA samples for mutational screening were obtained from peripheral blood leukocytes because somatic mutations are not expected to accumulate with aging in blood.³³ The study was approved by the Concord Hospital Human Research Ethics Committee, and all participants gave informed consent.

Total DNA was extracted using standard molecular procedures. A single 4.3-kb DNA fragment encompassing the three mitochondrial genes *MT-CO1*, *MT-CO2* and *MT-CO3* was polymerase chain reaction amplified from 100 CHAMP individuals using oligonucleotide primers MT-CO1F (5'-AGGTTTG AAGCTGCTTCTTCG-3'; positions 5771–5791) and MT-CO3R (5'-ATTAGTA GTAAGGCTAGGAGG-3'; positions 10 111–10 091). Primer positions are based on the revised Cambridge Reference Sequence (GenBank accession number NC_012920) and each primer sequence was analyzed by Primer-BLAST (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>) to ensure product specificity. A 100 ng of DNA using the Crimson LongAmp *Taq* DNA Polymerase (New England Biolabs, Ipswich, MA, USA) was amplified using a manual hot start of 98 °C for 2 min, followed by 95 °C for 3 min, 40 cycles of 95 °C for 15 s, 56 °C for 1 min and 65 °C for 3.5 min. For the last 30 cycles, the elongation step at 65 °C was increased by 5 s per cycle. This was followed by a further incubation at 68 °C for 10 min. Polymerase chain reaction fragments were purified using ExoSap (USB Amersham, Buckinghamshire, UK) before sequencing.

To identify mtDNA sequence variation, each 4.3-kb polymerase chain reaction fragment was directly sequenced using nested primers specific for *MT-CO1* (Seq1, 5'-TTCAATATGAAAATCACCTC-3', positions 5801–5820; Seq2, 5'-ACCTAACCATCTTCTCCTT-3', positions 6334–6352; Seq3, 5'-GATT CATCTTTCTTTTCACC-3', positions 6931–6950), *MT-CO2* (Seq4, 5'-AACCA TTTCATAACTTTGTC-3', positions 7531–7550; Seq5, 5'-AATAATTACATCAC AAGACG-3', positions 8041–8060) and *MT-CO3* (Seq6, 5'-GAAATCGTGTC GCCTTA-3', positions 9133–9150) to cover each gene region. Polymerase chain reaction templates were denatured at 96 °C for 1 min followed by primer extension in the presence of Big Dye Terminators v.3.1 (Applied Biosystems, Mulgrave, VIC, Australia) for 30 cycles of 96 °C for 15 s, 50 °C for 15 s and 60 °C for 4 min. Trace sequences were analyzed using the Sequencher 5.0 gene analysis software (Gene Codes, Ann Arbor, MI, USA).

RESULTS

Quaternary 3-D model

The top scoring human COX structure was structurally aligned with the original bovine COX (PBD: 3ASO) and sequence identities were extracted from the resulting subunit pairings using UCSF Chimera 1.6 (Supplementary Table 2).³⁴ The initial model also included the hemes and other metal centers. Evaluation of the initial model showed a significant number of structural clashes as determined by van der Waals overlap >0.6 Å. The initial model had 363 structural clashes compared with 27 in the original bovine COX.

To refine the initial structure, we used 200 steps of steepest decent and 100 steps of conjugate gradient minimization, resulting in 26 and 4 clashes, respectively. Structure refinement also improved protein model validity scores. One metric for reliability was the Molprobit score, which was 2.07 for the original bovine structure, 3.16 for the initial model, 2.33 for the steepest decent outcome and 2.05 for the conjugate gradient final model.³⁵ In addition, each of the human COX structures was evaluated with the protein validation tools Molprobit, ERRAT2, PROCHECK and WHATCHECK, and compared with the bovine COX crystal structure.^{35–38}

Structural parameters measured for functional prediction

The impact of mutation on COX function was inferred from a stringent sequence conservation evaluation and from three unique structural parameters. Only mutations containing several of the structurally measured perturbations were inferred to have a functional impact on human COX.

The sequence identity between 20 species across all 13 subunits ranged from 72% for subunit MT-CO2 to 91% for subunit MT-CO1 (Supplementary Table 2). Although the MT-CO2 subunit sequence identity was only 72%, the sequence similarity based on Blotsum62 similarity groups was 87%, indicating that an unambiguous alignment could be achieved for reliable structure generation. In addition, no gaps were present in the alignment of any subunit pairing; thus, no underdetermined unstructured regions were introduced into the human COX homology model. In a large multidomain protein like COX, a conservative change of isoleucine to valine (that is, the loss of a single methyl group) would not be expected to impact negatively on protein function. This statement remains true unless several other factors are involved, such as direct perturbation of a prosthetic group or a local rearrangement leading to extensive repacking and loss of favorable distant interactions.

The structure-based parameters included local volume or steric perturbation, subunit–subunit interaction interface perturbation and non-covalent energy of interaction loss or incompatibility. A qualitative assessment of side-chain atom addition or subtraction was interpreted from local packing density, potential for significant propagation of steric rearrangement (necessary rearrangement of many other amino acids) and the potential impact on hemes and other metal centers. For example, rather subtle and conservative mutations like L196I (in MT-CO1) involved in tight helix–helix packing, V38I (in MT-CO2) making direct contact with a heme prosthetic group (Table 1) and I175V (in MT-CO2) indirectly influencing the Cu_B ligands (Supplementary Table 3) are all predicted to have a significant impact on human COX function. However, surface-exposed mutations, either facing the aqueous environment or the lipid bilayer, were interpreted to have little or no impact on COX function. The lipid-exposed mutations included V91A and F251L (in MT-CO3) (Table 1) and V193I and I419V (in MT-CO1) (Supplementary Table 3), with each mutation being inferred to be inconsequential to COX function. The unique volume and steric properties of glycine and proline were also considered. For example, glycines in very tight turns when mutated could introduce steric limitations, like G125D (in MT-CO1), or interfere with very close helix–helix packing, like G170S (in MT-CO3) (Table 1). Likewise, steric limitations introduced with proline mutation can influence turn and secondary structure conformations. Proline mutants L135P (in MT-CO2) and S195P (in MT-CO3) are examples and were deemed significant to COX function (Table 1).

Mutations causing a reduction of subunit–subunit interaction energy could presumably reduce COX levels in the mitochondrial membrane. Mutations located in transmembrane helices are predicted to cause the greatest structural perturbations because of changes in packing density and regularity of the backbone structure. Any time a mutation resulted in a significant cross-subunit (subunit–subunit) interaction in the 3-D model, we assumed there would be a significant impact on protein assembly, stability or metal center assembly and environment. For example, the MT-CO1 mutation, F94C (Supplementary Table 3), makes significant hydrophobic contacts with MT-CO3 within the membrane spanning part of this interface. The loss of volume and hydrophobic surface with the mutation could reduce the interaction energy between the two subunits. The MT-CO2 mutation, K98Q (Table 1), makes extensive contact with COX6B1 on the matrix side of the membrane as well as an internal salt bridge that would be broken because of the mutation. Both the loss of a counter charge to a nearby glutamic acid and the change in side-chain volume would be expected to weaken the interaction between MT-CO2 and COX6B1. The MT-CO3 mutation, F35S (Supplementary Table 3),

Table 1 Quaternary modeling of disease-associated COX amino-acid variants previously tested for biochemical defect

Disease	Gene	Mutation	Steric prediction	Energy prediction	Subunit contact	Conservation	Total score	Modeller prediction	MutPred prediction ^a	Biochemical defect (reference)
<i>Rare inherited mitochondrial diseases</i>										
AH	MT-CO2	A41T	0	0	1	0 (15%)	1	0.25	0.506 (AH)	None ³⁹
E	MT-CO2	E129X	1	1	1	1 (100%)	4	1	—	Reduced COX ¹⁰
LD	MT-CO3	V91A	0.5	0	0	0.5 (90%)	1	0.25	0.539	None ⁴⁰
LD	MT-CO3	Q111>113X	1	1	1	1 (100%)	4	1	—	Reduced COX ⁴¹
LHON	MT-CO1	X514K	0	1	1	1 (100%)	3	0.75	—	Reduced COX ⁴²
LHON	MT-CO2	T13I	1	1	0.5	0.5 (75%)	3	0.75	0.396	Reduced respiration ⁴³
LHON	MT-CO3	M152L	0	0.5	0	1 (100%)	1.5	0.375	0.63	Reduced respiration ⁴³
MELAS	MT-CO1	V374M	1	1	0	1 (100%)	3	0.75	0.692 (AH)	Reduced COX ¹²
MELAS	MT-CO3	F251L	0	0	0	1 (95%)	1	0.25	0.686	None ⁴⁴
TE	MT-CO1	L196I	0.5	0	1	1 (100%)	2.5	0.625	0.698	Reduced COX ⁴⁵
<i>Mitochondrial-associated diseases</i>										
AD	MT-CO3	F219L	1	0.5	0	0 (50%)	1.5	0.375	0.461	Reduced COX ⁴⁶
AISA	MT-CO1	M273T	1	1	0	1 (100%)	3	0.75	0.718	Reduced COX ⁴⁷
AISA	MT-CO1	I280T	1	1	0	1 (100%)	3	0.75	0.766 (AH)	Reduced COX ⁴⁷
CC	MT-CO1	A32T	0.5	1	1	1 (100%)	3.5	0.875	0.591	Reduced COX ⁴⁸
CC	MT-CO1	G125D	1	1	1	1 (100%)	4	1	0.876	Reduced COX ⁴⁹
CC	MT-CO1	S458P	1	1	1	1 (100%)	4	1	0.796 (CH)	Reduced COX ⁴⁹
CC	MT-CO3	S14N	0.5	0.5	1	1 (100%)	3	0.75	0.732 (AH)	Reduced COX ⁴⁸
CC	MT-CO3	G170S	1	1	0	1 (100%)	3	0.75	0.618 (AH)	Reduced COX ⁴⁸
EXIT	MT-CO1	S142F	1	1	0	1 (100%)	3	0.75	0.825 (CH)	Reduced COX ⁵⁰
EXIT	MT-CO3	I256L	1	0	0	0.5 (90%)	1.5	0.375	0.706	Reduced COX ⁵¹
EXIT/MR	MT-CO1	G351D	1	1	0	1 (100%)	3	0.75	0.85	Reduced COX ⁵²
HC	MT-CO2	V38I	1	0	1	1 (100%)	3	0.75	0.6 (AH)	Reduced COX ⁵³
IE	COX6B1	R20H	0.5	1	0	1 (100%)	2.5	0.625	0.967	Reduced COX ⁵⁴
LA	MT-CO2	M153>155X	1	1	1	1 (100%)	4	1	—	Reduced COX ⁵⁵
M	MT-CO1	K265>271X	1	1	1	1 (100%)	4	1	—	Reduced COX ⁵⁶
M	MT-CO3	S195P	1	1	1	1 (95%)	4	1	0.756 (AH)	Reduced COX ¹⁰
MD	MT-CO1	G343X	1	1	1	1 (100%)	4	1	—	Reduced COX ⁵⁷
MD	MT-CO2	W104X	1	1	1	1 (100%)	4	1	—	Reduced COX ⁵⁸
ME	MT-CO2	M1T	1	1	1	1 (100%)	4	1	0.597 (AH)	Reduced COX ⁸
ME	MT-CO3	W249X	1	1	0	1 (100%)	3	0.75	—	Reduced COX ⁵⁹
Mg	MT-CO1	W6X	1	1	1	1 (100%)	4	1	—	Reduced COX ⁶⁰
Mg	MT-CO3	del-FFFAG91	1	1	1	1 (100%)	4	1	—	Reduced COX ⁶¹
MM	MT-CO2	M29K	0.5	1	1	1 (100%)	3.5	0.875	0.782 (CH)	Reduced COX ⁹
MM/LA	MT-CO3	W58X	1	1	1	1 (100%)	4	1	—	Reduced COX ⁶²
MM/R	MT-CO1	G269X	1	1	1	1 (100%)	4	1	—	Reduced COX ⁶³
MND	MT-CO1	delAELGQ39>45X	1	1	1	1 (100%)	4	1	—	Reduced COX ⁷
PEG	MT-CO2	K98Q	1	1	1	1 (100%)	4	1	0.746 (AH)	None ⁶⁴
PI	COX4I2	E138K	0	0.5	0	0.5 (90%)	1	0.25	0.906 (VCH)	Reduced mRNA ⁶⁵
R	MT-CO2	L135P	1	1	1	1 (100%)	4	1	0.909 (VCH)	Reduced COX ⁶⁶
R	MT-CO3	P118>124X	1	1	1	1 (100%)	4	1	—	Reduced COX ⁶⁷

Abbreviations: AD, Alzheimer's disease; AH, Alpers–Huttenlocher; AISA, acquired idiopathic sideroplatic anemia; CC, colon cancer; E, encephalopathy; EXIT, exercise intolerance; HC, hypertrophic cardiomyopathy; IE, infantile encephalomyopathy; LA, lactic acidosis; LD, Leigh disease; LHON, Leber hereditary optic neuropathy; M, myopathy; MD, multisystem disorder; ME, mitochondrial encephalomyopathy; MELAS, mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes; Mg, myoglobinuria; MM, mitochondrial myopathy; MND, motor neuron disease; MR, mental retardation; PEG, pseudoexfoliation glaucoma; PI, pancreatic insufficiency; R, rhabdomyolysis; TE, therapy-resistant epilepsy.

^aOnly variants with AH (actionable hypothesis, may have influence on disease), CH (confident hypothesis, confidence that variant is disease causing) and VCH (very confident hypothesis, high confidence level that variant is disease causing) represent significance.

makes extensive hydrophobic contacts with both MT-CO1 and COX6A1 near the membrane/solvent interface. The dramatic change to serine would change the volume greatly as well as the local polarity, both of which would be expected to reduce the stability of the subunit interactions.

Non-covalent interaction perturbations were inferred when polar residues or those having charged side chains occurred in an otherwise hydrophobic environment (for example, I280T in MT-CO1 and M1T

or M29K in MT-CO2) or when mutations caused disruption of an interacting polar or charged side chains (for example, S142F in MT-CO1, K98Q in MT-CO2 and S14N or S195P in MT-CO3) (Table 1). A perturbation was also assigned when rearrangement of the local side-chain was inferred to disrupt local hydrogen bonding or electrostatic interactions. However, these were assumed to have only a small impact on COX function (for example, F219L in MT-CO3) (Table 1).

COX variants listed in the human mutation database

Currently, 162 COX amino-acid mutations are listed in Mitomap.²⁷ Of these, 40 have known biochemical functionality (36 are reported with a biochemical defect, while four variants have normal function) and 122 are not biochemically characterized (Figure 1).

Of the 36 variants with reduced function, 34 are in mtDNA core COX subunits (MT-CO1 = 15; MT-CO2 = 8; MT-CO3 = 11) and two are in nuclear-encoded subunits (COX4I2 = 1; COX6B1 = 1). In total, 3-D model analysis of all 36 mutations predicted that 89% (32/36) were likely to be functionally important in a nuclear background that does not harbor compensatory mutations (Table 1). MutPred predictions estimated that 33% (12/36) of the disease-associated mutations with a reported biochemical defect would be functional.

3-D model analysis of the four variants with reported normal activity (A41T, K98Q in MT-CO2, and V91A, F251L in MT-CO3) supported the biochemical data except for the MT-CO2 K98Q variant, which produced a false-positive result (Table 1). MutPred analysis of these same four variants produced false-positive data for K98Q (in MT-CO2) and V91A (in MT-CO3) (Table 1).

A total of 122 non-functionally characterized core COX amino-acid mutations in Mitomap²⁷ were mapped onto the 3-D model (MT-CO1 = 52; MT-CO2 = 31; MT-CO3 = 39). The model predicts that 41% (50/122) core COX variants may be functionally important in a nuclear background without compensatory mutations (Supplementary Table 3). MutPred analysis of the non-functionally characterized core COX mutations estimated that 17% (21/122) would be functional (Supplementary Table 3).

Assaying COX variation from an aged cohort

We sequenced the three core COX genes from 100 aged men to determine the prevalence of mtDNA mutations and estimate the number of functional changes determined by the 3-D model and by MutPred.³² Sequence analysis revealed a total of 69 mtDNA mutations (MT-CO1 = 38; MT-CO2 = 17; MT-CO3 = 14). Of these, 57 (82.6%) were synonymous and 12 (17.4%) were amino-acid variants. From the 12 mutations that caused an amino-acid change,

eight are reported in Mitomap.²⁷ Of these eight, three are associated with prostate cancer, while five come from direct submissions and have unknown function (Table 2). The remaining four amino-acid changes are potentially novel.

The 3-D model predictions suggested that four of the 12 mutations (found in four of 100 individuals) were likely to alter the structural integrity of the COX complex and were functionally important (Table 2), assuming that compensatory nuclear mutations have also not occurred. Of the four variants predicted to be functional, two of these, A120T (in MT-CO1) and V136L (in MT-CO3), were predicted to interact with the nuclear-encoded subunits of COX7C and COX6A1, respectively, while the S330G variant (in MT-CO1) was predicted to interact with the mitochondrial-encoded MT-CO2 subunit. The remaining V143M variant (in MT-CO1) was not predicted to interact with any other subunit. MutPred predicted that one (MT-CO1, V193I) of these 12 amino-acid variants would be functional.

DISCUSSION

Here we develop a quaternary 3-D model for human mitochondrial COX. A strength of the model is that it can interpret and predict the structural consequences of amino-acid variation in all 13 protein subunits. Importantly, the influence of compensatory changes can also be modeled. In this study, we compare the results from the 3-D model against the MutPred web application tool.²⁶ However, direct comparisons between the 3-D model and MutPred need to be made with caution as the 3-D model requires an in-depth understanding of bioinformatics modeling. In contrast, MutPred is a fully automated web-based tool that can quickly analyze both structural and functional properties of each protein independently.

First, we investigated mutations reported in Mitomap. The 3-D model predicted that 89% (32/36) of variants previously reported with a measurable biochemical defect would be functionally important. In contrast, MutPred predicted that 33% (12/36) of these same variants would be functionally important. For non-biochemically characterized variants, the 3-D model predicted that 41% (50/122)

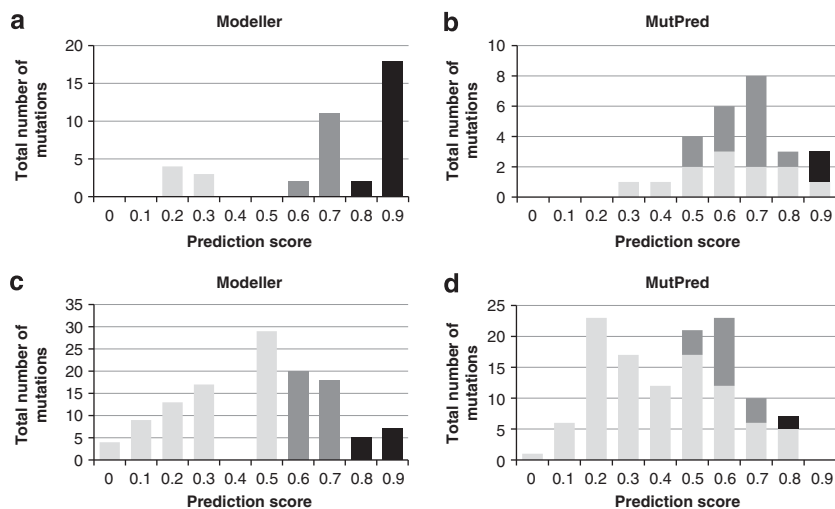


Figure 1 Comparisons between Modeller and MutPred for functional mutation prediction. **a** and **b** represent the total number of COX mutations with reported biochemical defects analyzed for functional prediction. **c** and **d** represent the total number of COX mutations with unknown functional activity analyzed for functional prediction. Light gray bars, the number of mutations predicted to be non-functional; dark gray bars, the number of mutations assigned a confidence level of being functional; and black bars, the number of mutations assigned a very confident level of being functional.

Table 2 Amino-acid variation identified from sequencing 100 CHAMP individuals

CHAMP sample	Gene	Mutation	Predicted function ^a	Associated with disease	Reference ^b
1	MT-CO2	T167S	Non-functional	—	This study
2	MT-CO1	D4N	Non-functional	Prostate	Petros <i>et al.</i> ¹⁴
3	MT-CO3	V91I	Non-functional	—	Online
4 ^c	MT-CO1	N46S	Non-functional	—	Online
4 ^c	MT-CO3	V91I	Non-functional	—	Online
5	MT-CO3	V91I	Non-functional	—	Online
6	MT-CO1	D4N	Non-functional	Prostate	Petros <i>et al.</i> ¹⁴
7	MT-CO3	V91I	Non-functional	—	Online
8	MT-CO3	V91I	Non-functional	—	Online
9	MT-CO3	A32T	Non-functional	—	Online
10	MT-CO2	V90I	Non-functional	—	Online
11	MT-CO1	D4N	Non-functional	Prostate	Petros <i>et al.</i> ¹⁴
12	MT-CO3	V91I	Non-functional	—	Online
13 ^c	MT-CO1	A120T	Functional	Prostate	Petros <i>et al.</i> ¹⁴
13 ^c	MT-CO1	S330G	Functional	—	Online
14	MT-CO3	V91I	Non-functional	—	Online
15	MT-CO3	V91I	Non-functional	—	Online
16	MT-CO3	V91I	Non-functional	—	Online
17	MT-CO3	V91I	Non-functional	—	Online
18	MT-CO3	V136L	Functional	—	This study
19 ^c	MT-CO1	V143M	Functional	—	This study
19 ^c	MT-CO1	V193I	Non-functional	Prostate	Petros <i>et al.</i> ¹⁴
20	MT-CO3	V91I	Non-functional	—	Online
21 ^c	MT-CO1	A120T	Functional	Prostate	Petros <i>et al.</i> ¹⁴
21 ^c	MT-CO1	A477S	Non-functional	—	This study

^aPredicted function from quaternary modeling.^bOnline in Mitomap.²⁷^cIndividuals carrying linked variants.

would influence COX function, while MutPred predicted 17% (21/122) would be functional (Supplementary Table 3). The proportion of mutations that is predicted to be deleterious by both the 3-D model and by MutPred suggests that the data present in Mitomap cannot be used to infer the frequency of deleterious mutations in a normal population. Elliott *et al.*⁶⁸ looked for 10 well-known rare inherited mitochondrial diseases, and found that 1/200 harbored them at very low levels of heteroplasmy, while Schaefer *et al.*⁶⁹ found that the frequency of mtDNA mutations in disease occurs in approximately 1/4000 individuals.

To test for the frequency of mutations predicted to be functional from within a known population, we isolated DNA and sequenced the three core COX genes from 100 subjects randomly selected from a cohort of 1705 aged individuals enrolled on the Sydney CHAMP study.³² Sequencing of DNA from older men (>70 years) indicates that about 20% (21/100) of the male aged population contain a COX amino-acid mutation. In total, 12 amino-acid variants were found in the aged cohort, with 4 of these predicted to be functional by the 3-D model (Table 2). In contrast, MutPred predicted that one of these 12 amino-acid variants would be functional. In combination with the results obtained from examining the mutations listed in Mitomap, these data show that the 3-D model predicts more functionally significant mutations than does MutPred. At this time, it seems likely that either the 3-D model is overpredicting functional

mutations and/or MutPred is underpredicting them. In this study, mutations generating normalized ratio scores of 0.6 and above were considered functionally important. This score may need to be modified as additional functional data on the significance of specific mutations becomes known.

There are several limitations to the current 3-D model. The model is incapable of analyzing some functional properties of protein residues (i.e., DNA-binding residues, phosphorylation and methylation sites). Furthermore, structural analysis on the locations of putative proton/water channels, specific binding surfaces (i.e., the cytochrome *c* binding surface) or specific lipid binding sites (lipids that could be involved in protein activity) were not measured. In this initial work, we decided not to specify amino acids directly involved in proton translocation through the COX D and K proton channels. These channels have been best described in bacterial COX mutagenesis and crystallographic studies.⁷⁰ The current model and, in fact, the bovine COX template do not contain the crystallographically resolved water molecules comprising these proton channels. Some of the impact of mutating proton channel residues would be accounted for in the current analysis as energetic perturbations in highly conserved positions. With respect to the cytochrome *c* binding surface, the exact location(s) of interaction has not yet been reported nor has the impact of individual amino-acid changes. Consideration could be given to clusters of aspartate and/or glutamate in future models as cytochrome *c* itself contains positive amino acids on its surface surrounding its heme. Perhaps, the most difficult functional interaction to characterize is with structurally/functionally important lipids. For example, COX function is known to depend on cardiolipin, but the location of these and other lipids vary or do not exist in current crystallographic structures. This is a likely source of error in predicting mutational impact on the unusually poorly packed MT-CO3, which has significant lipid exposure. In the amino-acid conservation analysis, a fairly stringent metric was used and no consideration was given to similar substitutions (i.e., glutamic acid to aspartic acid). Finally, there is also some potential for double counting importance between volume and energy as they are not completely independent.

This study strongly suggests that the 3-D model will be a useful tool for the functional determination of specific COX mutations. An advantage of the 3-D model is that the loss of function criteria can be adjusted and that insight can be gained about the presence of compensatory mechanisms. Additional sampling of mitochondrial and nuclear-encoded genes in young and older individuals along with the functional testing of the bioenergetic consequences of specific changes is however required to fully assess the robustness of the model and the importance of compensatory mechanisms in humans. Furthermore, this study may form the basis of future studies to generate 3-D models for each of the mitochondrial electron transport system complexes.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank the investigators involved with the CHAMP study, including Professor Robert Cumming, Associate Professor Fiona Blyth, Associate Professor Vasikaran Naganathan, Professor David Handelsman and Melisa Litchfield for the kind provision of DNA and epidemiological data. We also thank members of the Ballard lab group for comments on the manuscript.

- 1 Shoubridge, E. A. Cytochrome *c* oxidase deficiency. *Am. J. Med. Genet.* **106**, 46–52 (2001).
- 2 Villani, G., Greco, M., Papa, S. & Attardi, G. Low reserve of cytochrome *c* oxidase capacity *in vivo* in the respiratory chain of a variety of human cell types. *J. Biol. Chem.* **273**, 31829–31836 (1998).
- 3 Barrientos, A., Barros, M. H., Valnot, I., Rotig, A., Rustin, P. & Tzagoloff, A. Cytochrome oxidase in health and disease. *Gene* **286**, 53–63 (2002).
- 4 Ongwijitwat, S. & Wong-Riley, M. T. Is nuclear respiratory factor 2 a master transcriptional coordinator for all ten nuclear-encoded cytochrome *c* oxidase subunits in neurons? *Gene* **360**, 65–77 (2005).
- 5 Dhar, S. S., Ongwijitwat, S. & Wong-Riley, M. T. Nuclear respiratory factor 1 regulates all ten nuclear-encoded subunits of cytochrome *c* oxidase in neurons. *J. Biol. Chem.* **283**, 3120–3129 (2008).
- 6 Osada, N. & Akashi, H. Mitochondrial–nuclear interactions and accelerated compensatory evolution: evidence from the primate cytochrome *C* oxidase complex. *Mol. Biol. Evol.* **29**, 337–346 (2012).
- 7 Comi, G. P., Bordoni, A., Salani, S., Franceschina, L., Sciacco, M., Prella, A. *et al.* Cytochrome *c* oxidase subunit I microdeletion in a patient with motor neuron disease. *Ann. Neurol.* **43**, 110–116 (1998).
- 8 Clark, K. M., Taylor, R. W., Johnson, M. A., Chinnery, P. F., Chrzanoska-Lightowlers, Z. M., Andrews, R. M. *et al.* An mtDNA mutation in the initiation codon of the cytochrome *C* oxidase subunit II gene results in lower levels of the protein and a mitochondrial encephalomyopathy. *Am. J. Hum. Genet.* **64**, 1330–1339 (1999).
- 9 Rahman, S., Taanman, J. W., Cooper, J. M., Nelson, I., Hargreaves, I., Meunier, B. *et al.* A missense mutation of cytochrome oxidase subunit II causes defective assembly and myopathy. *Am. J. Hum. Genet.* **65**, 1030–1039 (1999).
- 10 Horvath, R., Schoser, B. G., Muller-Hockner, J., Volpel, M., Jaksch, M. & Lochmuller, H. Mutations in mtDNA-encoded cytochrome *c* oxidase subunit genes causing isolated myopathy or severe encephalomyopathy. *Neuromuscul. Disord.* **15**, 851–857 (2005).
- 11 Hoffbuhr, K. C., Davidson, E., Filiano, B. A., Davidson, M., Kennaway, N. G. & King, M. P. A pathogenic 15-base pair deletion in mitochondrial DNA-encoded cytochrome *c* oxidase subunit III results in the absence of functional cytochrome *c* oxidase. *J. Biol. Chem.* **275**, 13994–14003 (2000).
- 12 Tam, E. W., Feigenbaum, A., Addis, J. B., Blaser, S., Mackay, N., Al-Dosary, M. *et al.* A novel mitochondrial DNA mutation in COX1 leads to strokes, seizures, and lactic acidosis. *Neuropediatrics* **39**, 328–334 (2008).
- 13 Bratic, I. & Trifunovic, A. Mitochondrial energy metabolism and ageing. *Biochim. Biophys. Acta* **1797**, 961–967 (2010).
- 14 Petros, J. A., Baumann, A. K., Ruiz-Pesini, E., Amin, M. B., Sun, C. Q., Hall, J. *et al.* mtDNA mutations increase tumorigenicity in prostate cancer. *Proc. Natl Acad. Sci. USA* **102**, 719–724 (2005).
- 15 Diaz, F. Cytochrome *c* oxidase deficiency: patients and animal models. *Biochim. Biophys. Acta* **1802**, 100–110 (2010).
- 16 Raha, S., Merante, F., Shoubridge, E., Myint, A. T., Tein, I., Benson, L. *et al.* Repopulation of rhoO cells with mitochondria from a patient with a mitochondrial DNA point mutation in tRNA(Gly) results in respiratory chain dysfunction. *Hum. Mutat.* **13**, 245–254 (1999).
- 17 Baranowska, I., Jaderlund, K. H., Nennesmo, I., Holmqvist, E., Heidrich, N., Larsson, N. G. *et al.* Sensory ataxic neuropathy in golden retriever dogs is caused by a deletion in the mitochondrial tRNATyr gene. *PLoS Genet.* **5**, e1000499 (2009).
- 18 Hiona, A., Sanz, A., Kujoth, G. C., Pamplona, R., Seo, A. Y., Hofer, T. *et al.* Mitochondrial DNA mutations induce mitochondrial dysfunction, apoptosis and sarcopenia in skeletal muscle of mitochondrial DNA mutator mice. *PLoS One* **5**, e11468 (2010).
- 19 Dantzer, J., Moad, C., Heiland, R. & Mooney, S. MutDB services: interactive structural analysis of mutation data. *Nucleic Acids Res.* **33**, W311–W314 (2005).
- 20 Pinto, J. R., Parvatiyar, M. S., Jones, M. A., Liang, J., Ackerman, M. J. & Potter, J. D. A functional and structural study of troponin C mutations related to hypertrophic cardiomyopathy. *J. Biol. Chem.* **284**, 19090–19100 (2009).
- 21 Hicks, S., Wheeler, D. A., Plon, S. E. & Kimmel, M. Prediction of missense mutation functionality depends on both the algorithm and sequence alignment employed. *Hum. Mutat.* **32**, 661–668 (2011).
- 22 Melvin, R. G., Katewa, S. D. & Ballard, J. W. O. A candidate complex approach to study functional mitochondrial DNA changes: sequence variation and quaternary structure modeling of *Drosophila simulans* cytochrome *c* oxidase. *J. Mol. Evol.* **66**, 232–242 (2008).
- 23 Tsukihara, T., Aoyama, H., Yamashita, E., Tomizaki, T., Yamaguchi, H., Shinzawa-Itoh, K. *et al.* Structures of metal sites of oxidized bovine heart cytochrome *c* oxidase at 2.8 Å. *Science* **269**, 1069–1074 (1995).
- 24 Suga, M., Yano, N., Muramoto, K., Shinzawa-Itoh, K., Maeda, T., Yamashita, E. *et al.* Distinguishing between Cl[−] and O₂^{2−} as the bridging element between Fe³⁺ and Cu²⁺ in resting-oxidized cytochrome *c* oxidase. *Acta Crystallogr. D* **67**, 742–744 (2011).
- 25 Sali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
- 26 Li, B., Krishnan, V. G., Mort, M. E., Xin, F., Kamati, K. K., Cooper, D. N. *et al.* Automated inference of molecular mechanisms of disease from amino acid substitutions. *Bioinformatics* **25**, 2744–2750 (2009).
- 27 Ruiz-Pesini, E., Lott, M. T., Procaccio, V., Poole, J. C., Brandon, M. C., Mishmar, D. *et al.* An enhanced MITOMAP with a global mtDNA mutational phylogeny. *Nucleic Acids Res.* **35**, D823–D828 (2007).
- 28 Pei, J. & Grishin, N. V. PROMALS: towards accurate multiple sequence alignments of distantly related proteins. *Bioinformatics* **23**, 802–808 (2007).
- 29 Reid, B. P. O. L. M. & Sparks, R. K. Methane–methane isotropic interaction potential from total differential cross sections. *J. Chem. Phys.* **11**, 5656 (1985).
- 30 Ratnaparkhi, G. S. & Varadarajan, R. Thermodynamic and structural studies of cavity formation in proteins suggest that loss of packing interactions rather than the hydrophobic effect dominates the observed energetics. *Biochemistry* **39**, 12365–12374 (2000).
- 31 Holder, J. B., Bennett, A. F., Chen, J., Spencer, D. S., Byrne, M. P. & Stites, W. E. Energetics of side chain packing in staphylococcal nuclease assessed by exchange of valines, isoleucines, and leucines. *Biochemistry* **40**, 13998–14003 (2001).
- 32 Cumming, R. G., Handelsman, D., Seibel, M. J., Creasey, H., Sambrook, P., Waite, L. *et al.* Cohort Profile: the Concord Health and Ageing in Men Project (CHAMP). *Int. J. Epidemiol.* **38**, 374–378 (2009).
- 33 t Hart, L. M., Jansen, J. J., Lemkes, H. H., de Knijff, P. & Maassen, J. A. Heteroplasmy levels of a mitochondrial gene mutation associated with diabetes mellitus decrease in leucocyte DNA upon aging. *Hum. Mutat.* **7**, 193–197 (1996).
- 34 Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. *et al.* UCSF chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
- 35 Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X. *et al.* MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383 (2007).
- 36 Colovos, C. & Yeates, T. O. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci.* **2**, 1511–1519 (1993).
- 37 Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486 (1996).
- 38 Hooft, R. W., Vriend, G., Sander, C. & Abola, E. E. Errors in protein structures. *Nature* **381**, 272 (1996).
- 39 Uusimaa, J., Finnila, S., Vainionpaa, L., Karpaa, M., Herva, R., Rantala, H. *et al.* A mutation in mitochondrial DNA-encoded cytochrome *c* oxidase II gene in a child with Alpers–Huttenlocher-like disease. *Pediatrics* **111**, e262–e268 (2003).
- 40 Mkaouer-Rebai, E., Ellouze, E., Chamkha, I., Kammoun, F., Triki, C. & Fakhfakh, F. Molecular–clinical correlation in a family with a novel heteroplasmic Leigh syndrome missense mutation in the mitochondrial cytochrome *c* oxidase III gene. *J. Child Neurol.* **26**, 12–20 (2011).
- 41 Tiranti, V., Corona, P., Greco, M., Taanman, J. W., Carrara, F., Lamantea, E. *et al.* A novel frameshift mutation of the mtDNA COIII gene leads to impaired assembly of cytochrome *c* oxidase in a patient affected by Leigh-like syndrome. *Hum. Mol. Genet.* **9**, 2733–2742 (2000).
- 42 Brown, M. D., Yang, C. C., Trounce, I., Torroni, A., Lott, M. T. & Wallace, D. C. A mitochondrial DNA variant, identified in Leber hereditary optic neuropathy patients, which extends the amino acid sequence of cytochrome *c* oxidase subunit I. *Am. J. Hum. Genet.* **51**, 378–385 (1992).
- 43 Abu-Amero, K. K. & Bosley, T. M. Mitochondrial abnormalities in patients with LHON-like optic neuropathies. *Invest. Ophthalmol. Vis. Sci.* **47**, 4211–4220 (2006).
- 44 Manfredi, G., Schon, E. A., Moraes, C. T., Bonilla, E., Berry, G. T., Sladky, J. T. *et al.* A new mutation associated with MELAS is located in a mitochondrial DNA polypeptide-coding gene. *Neuromuscul. Disord.* **5**, 391–398 (1995).
- 45 Varlamov, D. A., Kudin, A. P., Vielhaber, S., Schroder, R., Sassen, R., Becker, A. *et al.* Metabolic consequences of a novel missense mutation of the mtDNA CO I gene. *Hum. Mol. Genet.* **11**, 1797–1805 (2002).
- 46 Hamblet, N. S., Ragland, B., Ali, M., Conyers, B. & Castora, F. J. Mutations in mitochondrial-encoded cytochrome *c* oxidase subunits I, II, and III genes detected in Alzheimer's disease using single-strand conformation polymorphism. *Electrophoresis* **27**, 398–408 (2006).
- 47 Broker, S., Meunier, B., Rich, P., Gattermann, N. & Hofhaus, G. MtDNA mutations associated with sideroblastic anaemia cause a defect of mitochondrial cytochrome *c* oxidase. *Eur. J. Biochem.* **258**, 132–138 (1998).
- 48 Greaves, L. C., Barron, M. J., Plusa, S., Kirkwood, T. B., Mathers, J. C., Taylor, R. W. *et al.* Defects in multiple complexes of the respiratory chain are present in ageing human colon crypts. *Exp. Gerontol.* **45**, 573–579 (2010).
- 49 Namslauer, I. & Brzezinski, P. A mitochondrial DNA mutation linked to colon cancer results in proton leaks in cytochrome *c* oxidase. *Proc. Natl Acad. Sci. USA* **106**, 3402–3407 (2009).
- 50 Lucifoli, S., Hoffmeier, K., Carozzo, R., Tessa, A., Ludwig, B. & Santorelli, F. M. Introducing a novel human mtDNA mutation into the *Paracoccus denitrificans* COX I gene explains functional deficits in a patient. *Neurogenetics* **7**, 51–57 (2006).
- 51 Bortot, B., Barbi, E., Biffi, S., Angelini, C., Faleschini, E., Severini, G. M. *et al.* Two novel cosegregating mutations in tRNAMet and COX III, in a patient with exercise intolerance and autoimmune polyendocrinopathy. *Mitochondrion* **9**, 123–129 (2009).
- 52 Herrero-Martin, M. D., Pineda, M., Briones, P., Lopez-Gallardo, E., Carreras, M., Benac, M. *et al.* A new pathologic mitochondrial DNA mutation in the cytochrome oxidase subunit I (MT-CO1). *Hum. Mutat.* **29**, E112–E122 (2008).
- 53 Wei, Y. L., Yu, C. A., Yang, P., Li, A. L., Wen, J. Y., Zhao, S. M. *et al.* Novel mitochondrial DNA mutations associated with Chinese familial hypertrophic cardiomyopathy. *Clin. Exp. Pharmacol. Physiol.* **36**, 933–939 (2009).
- 54 Massa, V., Fernandez-Vizcarra, E., Alshahwan, S., Baksh, E., Goffrini, P., Ferrero, I. *et al.* Severe infantile encephalomyopathy caused by a mutation in COX6B1, a nucleus-encoded subunit of cytochrome *c* oxidase. *Am. J. Hum. Genet.* **82**, 1281–1289 (2008).
- 55 Wong, L. J., Dai, P., Tan, D., Lipson, M., Grix, A., Sifry-Platt, M. *et al.* Severe lactic acidosis caused by a novel frame-shift mutation in mitochondrial-encoded cytochrome *c* oxidase subunit II. *Am. J. Med. Genet.* **102**, 95–99 (2001).

- 56 Valente, L., Piga, D., Lamantea, E., Carrara, F., Uziel, G., Cudia, P. *et al.* Identification of novel mutations in five patients with mitochondrial encephalomyopathy. *Biochim. Biophys. Acta* **1787**, 491–501 (2009).
- 57 Bruno, C., Martinuzzi, A., Tang, Y., Andreu, A. L., Pallotti, F., Bonilla, E. *et al.* A stop-codon mutation in the human mtDNA cytochrome c oxidase I gene disrupts the functional structure of complex IV. *Am. J. Hum. Genet.* **65**, 611–620 (1999).
- 58 Campos, Y., Garcia-Redondo, A., Fernandez-Moreno, M. A., Martinez-Pardo, M., Goda, G., Rubio, J. C. *et al.* Early-onset multisystem mitochondrial disorder caused by a nonsense mutation in the mitochondrial DNA cytochrome C oxidase II gene. *Ann. Neurol.* **50**, 409–413 (2001).
- 59 Hanna, M. G., Nelson, I. P., Rahman, S., Lane, R. J., Land, J., Heales, S. *et al.* Cytochrome c oxidase deficiency associated with the first stop-codon point mutation in human mtDNA. *Am. J. Hum. Genet.* **63**, 29–36 (1998).
- 60 Karadimas, C. L., Greenstein, P., Sue, C. M., Joseph, J. T., Tanji, K., Haller, R. G. *et al.* Recurrent myoglobinuria due to a nonsense mutation in the COX I gene of mitochondrial DNA. *Neurology* **55**, 644–649 (2000).
- 61 Keightley, J. A., Hoffbuhr, K. C., Burton, M. D., Salas, V. M., Johnston, W. S., Penn, A. M. *et al.* A microdeletion in cytochrome c oxidase (COX) subunit III associated with COX deficiency and recurrent myoglobinuria. *Nat. Genet.* **12**, 410–416 (1996).
- 62 Horvath, R., Scharfe, C., Hoeltzenbein, M., Do, B. H., Schroder, C., Warzok, R. *et al.* Childhood onset mitochondrial myopathy and lactic acidosis caused by a stop mutation in the mitochondrial cytochrome c oxidase III gene. *J. Med. Genet.* **39**, 812–816 (2002).
- 63 Kollberg, G., Moslemi, A. R., Lindberg, C., Holme, E. & Oldfors, A. Mitochondrial myopathy and rhabdomyolysis associated with a novel nonsense mutation in the gene encoding cytochrome c oxidase subunit I. *J. Neuropathol. Exp. Neurol.* **64**, 123–128 (2005).
- 64 Abu-Amero, K. K., Bosley, T. M. & Morales, J. Analysis of nuclear and mitochondrial genes in patients with pseudoexfoliation glaucoma. *Mol. Vis.* **14**, 29–36 (2008).
- 65 Shteyer, E., Saada, A., Shaag, A., Al-Hijawi, F. A., Kidess, R., Revel-Vilk, S. *et al.* Exocrine pancreatic insufficiency, dyserythropoietic anemia, and calvarial hyperostosis are caused by a mutation in the COX4I2 gene. *Am. J. Hum. Genet.* **84**, 412–417 (2009).
- 66 McFarland, R., Taylor, R. W., Chinnery, P. F., Howell, N. & Turnbull, D. M. A novel sporadic mutation in cytochrome c oxidase subunit II as a cause of rhabdomyolysis. *Neuromuscul. Disord.* **14**, 162–166 (2004).
- 67 Marotta, R., Chin, J., Kirby, D. M., Chiotis, M., Cook, M. & Collins, S. J. Novel single base pair COX III subunit deletion of mitochondrial DNA associated with rhabdomyolysis. *J. Clin. Neurosci.* **18**, 290–292 (2011).
- 68 Elliott, H. R., Samuels, D. C., Eden, J. A., Relton, C. L. & Chinnery, P. F. Pathogenic mitochondrial DNA mutations are common in the general population. *Am. J. Hum. Genet.* **83**, 254–260 (2008).
- 69 Schaefer, A. M., McFarland, R., Blakely, E. L., He, L., Whittaker, R. G., Taylor, R. W. *et al.* Prevalence of mitochondrial DNA disease in adults. *Ann. Neurol.* **63**, 35–39 (2008).
- 70 Yoshikawa, S., Muramoto, K. & Shinzawa-Itoh, K. Proton-pumping mechanism of cytochrome C oxidase. *Annu. Rev. Biophys.* **40**, 205–223 (2011).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)