

Research article

Open Access

## Epistatic interactions modulate the evolution of mammalian mitochondrial respiratory complex components

Luísa Azevedo\*<sup>†1</sup>, João Carneiro<sup>†1,2</sup>, Barbara van Asch<sup>1,2</sup>, Ana Moleirinho<sup>1</sup>, Filipe Pereira<sup>1,2</sup> and António Amorim<sup>1,2</sup>

Address: <sup>1</sup>IPATIMUP-Institute of Molecular Pathology and Immunology of the University of Porto, Porto, Portugal and <sup>2</sup>Faculty of Sciences of the University of Porto, Porto, Portugal

Email: Luísa Azevedo\* - lazevedo@ipatimup.pt; João Carneiro - jcarneiro@ipatimup.pt; Barbara van Asch - basch@ipatimup.pt; Ana Moleirinho - amoleirinho@ipatimup.pt; Filipe Pereira - fpereira@ipatimup.pt; António Amorim - aamorim@ipatimup.pt

\* Corresponding author †Equal contributors

Published: 13 June 2009

Received: 30 December 2008

BMC Genomics 2009, 10:266 doi:10.1186/1471-2164-10-266

Accepted: 13 June 2009

This article is available from: <http://www.biomedcentral.com/1471-2164/10/266>

© 2009 Azevedo et al., licensee BioMed Central Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** The deleterious effect of a mutation can be reverted by a second-site interacting residue. This is an epistatic compensatory process explaining why mutations that are deleterious in some species are tolerated in phylogenetically related lineages, rendering evident that those mutations are, by all means, only deleterious in the species-specific context. Although an extensive and refined theoretical framework on compensatory evolution does exist, the supporting evidence remains limited, especially for protein models. In this current study, we focused on the molecular mechanism underlying the epistatic compensatory process in mammalian mitochondrial OXPHOS proteins using a combination of in-depth structural and sequence analyses.

**Results:** Modeled human structures were used in this study to predict the structural impairment and recovery of deleterious mutations alone and combined with an interacting compensatory partner, respectively. In two cases, COI and COIII, intramolecular interactions between spatially linked residues restore the folding pattern impaired by the deleterious mutation. In a third case, intermolecular contact between mitochondrial CYB and nuclear CYTI encoded components of the cytochrome *bc1* complex are likely to restore protein binding. Moreover, we observed different modes of compensatory evolution that have resulted in either a quasi-simultaneous occurrence of a mutation and corresponding compensatory partner, or in independent occurrences of mutations in distinct lineages that were always preceded by the compensatory site.

**Conclusion:** Epistatic interactions between individual replacements involving deleterious mutations seems to follow a parsimonious model of evolution in which genomes hold pre-compensating states that subsequently tolerate deleterious mutations. This phenomenon is likely to have been constraining the variability at coevolving sites and shaping the interaction between the mitochondrial and the nuclear genome.

## Background

The deleterious impact of mutations can be reverted by epistatic interaction with a second-site which acts as a compensatory partner through a process known as compensatory evolution [1]. A detrimental mutation coupled to its compensatory partner results in a compensated background. As long as the deleterious effect of a mutation is neutralized, they find a chance to persist until possibly reaching fixation.

Several examples of compensatory evolution have been reported [2-9], with the great majority of studies involving coevolving nucleotide pairs that participate in the structural folding of RNA molecules. However, theoretical and empirical data are scarce for protein models [10-16]. This remains the case despite the expected impact that a network of intramolecular and intermolecular compensatory epistatic interactions may have on the expression of phenotypes [17] and protein evolution.

Although the most uncomplicated model of compensatory interaction comprises a two-locus, two-allele system, each compensated background may hold more than a single compensatory site for each deleterious mutation [2,18]. In addition, the functional rescue of each particular mutation may involve distinct compensatory solutions [19,20]. A recent study [21], focusing on the structural and physico-chemical properties of compensated mutations, revealed that the probability of a residue being compensated depends upon its location in the protein structure and on the degree of similarity between the changed and the newly arisen residues. That is to say; structurally exposed mutations and replacements involving similar residues are more easily compensated. Overall, these positions seem to represent about 10% of the observed replacements in a protein [12,13,22].

One of the most interesting issues concerning the interaction between coevolving residues is the succession of events that precede the fixation of a compensated genetic background. Two distinct explaining models have been proposed to explain this issue. In one model, a deleterious mutation arises in a background already harboring a fixed or polymorphic compensatory solution [10,12,23]. Because this deleterious effect is suppressed from the moment it occurs, it appears that the occurrence of human deleterious mutations in other mammals [10,12,24] can easily follow this model. Recently, we have shown that a lethal mutation in humans is naturally compensated in chimpanzees. In this case, the ancestral allele acts as a compensatory residue by restoring the protein activity to levels equivalent to the wild-type background [14]. However, unambiguous evidence regarding the recurrence of this model in compensatory evolution still awaits large-scale analyses. In the alternative model, a del-

eterious mutation occurs first, being subsequently compensated by a second-site substitution. This model presents two possible scenarios. In the first scenario, a mutation persists in the population at low frequencies while waiting for a compensatory solution to arise *de novo* by mutation [25,26], or as a polymorphism that may become associated via recombination [10]. In the second scenario, the fixation of the mutation always precedes the occurrence/fixation of its compensatory partner [27,28]. As expected, this scenario would not apply to strong fitness-affecting mutations.

Herein, the availability of data for a large number of related species motivated the selection of mitochondrial-encoded (OXPHOS) proteins to elucidate the molecular mechanism by which a compensatory residue provides structural, and consequently functional, rescue for the damaging effect of a deleterious replacement through the use of modeled structures. In addition, taking advantage of the large amount of available data for mammalian species, we pursued an extensive phylogenetic analysis to illustrate the sequence of events by which a mutation and a compensatory partner became paired in a compensated background.

## Results and Discussion

### **Identification of human deleterious mutations in non-human mammals**

At the time of this study, information was available for a total of 199 mammalian species as well as for 49 missense disease-associated mutations occurring at mitochondrial-encoded OXPHOS proteins (detailed information is given in the methods section). We started the analysis by identifying human missense pathological mutations representing the wild-type residue at the homologous site in non-human mammals, following a previously reported methodology [12]. A total of three unambiguously deleterious mutations [29-31], one at cytochrome c oxidase subunit I (COI), another at cytochrome c oxidase subunit III (COIII), and a third at cytochrome b (CYB) were identified and further examined (Table 1).

The COI-Leu196Ile mutation was detected in a patient suffering from epilepsy partialis continua [31]. Comparison of protein sequences revealed that the Leu196 residue is invariant in mammals indicating a critical role for normal protein function. Nevertheless, this mutation overlaps the wild-type residue in a rodent lineage (*M. glis*). The COIII-Phe251Leu mutation, known to be associated with mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes (MELAS) [29] was observed in seven primate species, all members of the Cercopithecidae family (*C. aethiops*, *C. tanzanus*, *C. sabaues*, *C. pygerythrus*, *S. entellus*, *M. mulatta* and *M. sylvanus*). Finally, the CYB-Gly251Ser replacement was detected in a patient present-

**Table 1: Human deleterious mutations at mitochondrial-encoded proteins present as wild-type amino acid in mammalian orthologues.**

Protein	Human mutation (associated phenotype)	Non-human mutation carriers
COI (Cytochrome c oxidase subunit I)	Leu196Ile (Epilepsy)	Order: Rodentia Family: Gliridae Species: <i>Myoxus glis</i> (Fat dormouse)
COIII (Cytochrome c oxidase subunit III)	Phe251Leu (Encephalopathy/MELAS)	Order: Primates Family: Cercopithecidae Species: <i>Chlorocebus aethiops</i> (African green monkey); <i>Chlorocebus tantalus</i> (Tantulus monkey); <i>Chlorocebus sabaeus</i> (Green monkey); <i>Chlorocebus pygerythrus</i> (Vervet monkey); <i>Semnopithecus entellus</i> (Hanuman langur); <i>Macaca mulatta</i> (Rhesus monkey); <i>Macaca sylvanus</i> (Barbary ape)
CYB (Cytochrome b)	Gly251Ser (Exercise intolerance)	Order: Primates Family: Hominidae Species: <i>Pongo pygmaeus</i> (Bornean orangutan); <i>Pongo abelii</i> (Sumatran orangutan)  Order: Primates Family: Cercopithecidae Species: <i>Macaca sylvanus</i> (Barbary ape); <i>Colobus guereza</i> (Guereza); <i>Trachypithecus obscurus</i> (Dusky leaf monkey); <i>Rhinopithecus roxellana</i> (Golden snub-nosed monkey); <i>Presbytis melalophos</i> (Mitre leaf monkey); <i>Procolobus badius</i> (Western red colobus); <i>Semnopithecus entellus</i> (Hanuman langur)  Order: Diprotodontia Family: Phalangeridae Species: <i>Trichosurus vulpecula</i> (Silver-gray brushtail possum)  Order: Diprotodontia Family: Phascolarctidae Species: <i>Phascolarctos cinereus</i> (Koala)

ing paracrystalline inclusions and low aerobic capacity [30]. This deleterious Ser251 residue was found in nine primate species (*P. pygmaeus*, *P. abelii*, *M. sylvanus*, *C. guereza*, *T. obscurus*, *R. roxellana*, *P. melalophos*, *P. badius* and *S. entellus*) as well as in two marsupial species in the Diprotodontia order (*T. vulpecula* and *P. cinereus*).

#### Uncovering mutation-compensation pairs

We next concentrated on the identification of the most likely compensatory site under the previous assumption [12] which stated that if a mutation is deleterious in humans but neutral in related species, protein sequence comparisons should reveal a compensatory site that distinguishes human and non-human wild-type sequences. That is, a compensatory residue must be recognizable in non-human mutation carriers and simultaneously be absent from the normal human sequence.

Here we propose a further extension to these assumptions namely, if a mutation is as deleterious in non-human mammals as it is in humans unless paired with a compensatory partner, the corresponding compensatory residue should be able to rescue the impairment of the human

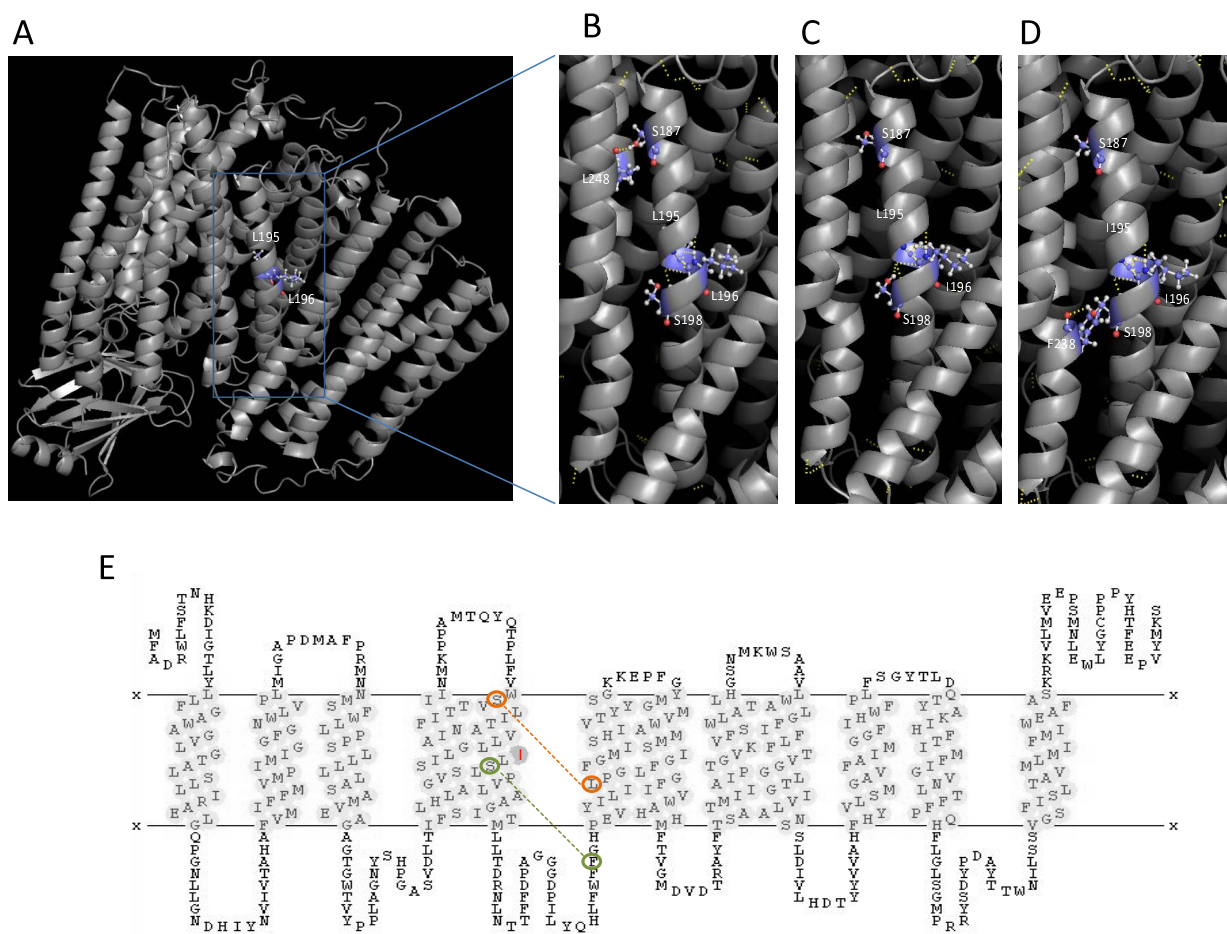
protein. In this perspective, the identification of the most likely compensatory partner for each mutation was initially based on comparative sequence data and subsequently complemented with 2D and 3D structural analyses to allow the recognition of the molecular basis of interactions. These analyses resulted in a vast list of potential compensatory sites for each mutation [see Additional file 1] from which the interacting residues in close spatial vicinity [12,15,16,18,32-34] were selected. Because the crystal structure of the human COI, COIII and CYB proteins has not been solved thus far, we built the structural models using previously established structures of bovine proteins [35,36] as templates (detailed information is given in the methods section).

Under these models, the compensatory residue for the deleterious COI-Ile196 found in *M. glis* was predicted to be at position 195 (Table 2) represented by an isoleucine in this lineage whereas all the other mammals preserve the leucine. Additional analyses provided compelling support for the compensatory role of Ile195 when interacting with deleterious Ile196. In the human modeled wild-type background (Leu195-Leu196, Figure 1A), side-chain

**Table 2: Candidate compensatory residues for three human deleterious mutations.**

Protein	Human mutation	Compensatory residues
COI	Leu196Ile	Ile195
COIII	Phe251Leu	Val254
CYB	Gly251Ser	Pro258, Ser263

interactions between  $\alpha$ -helix V and VI are achieved, at least partially, through an H-bond involving Ser187 ( $\alpha$ -helix V) and Leu248 ( $\alpha$ -helix VI) (Figure 1B and Figure 1E). This bonding was not predicted in the deleterious background (Leu195-Ile196, Figure 1C), strongly indicating an impairment in protein folding, a hypothesis substantiated by functional evaluations of the affected patient [31]. The connection between  $\alpha$ -helices V and VI was restored when the conserved Leu195 was replaced by an Ile, a combination that represent the compensated background (Ile195-Ile196, Figure 1D and Figure 1E). Two interesting points arise here. First, the novel interaction also involves novel intervening residues (Ser198 and



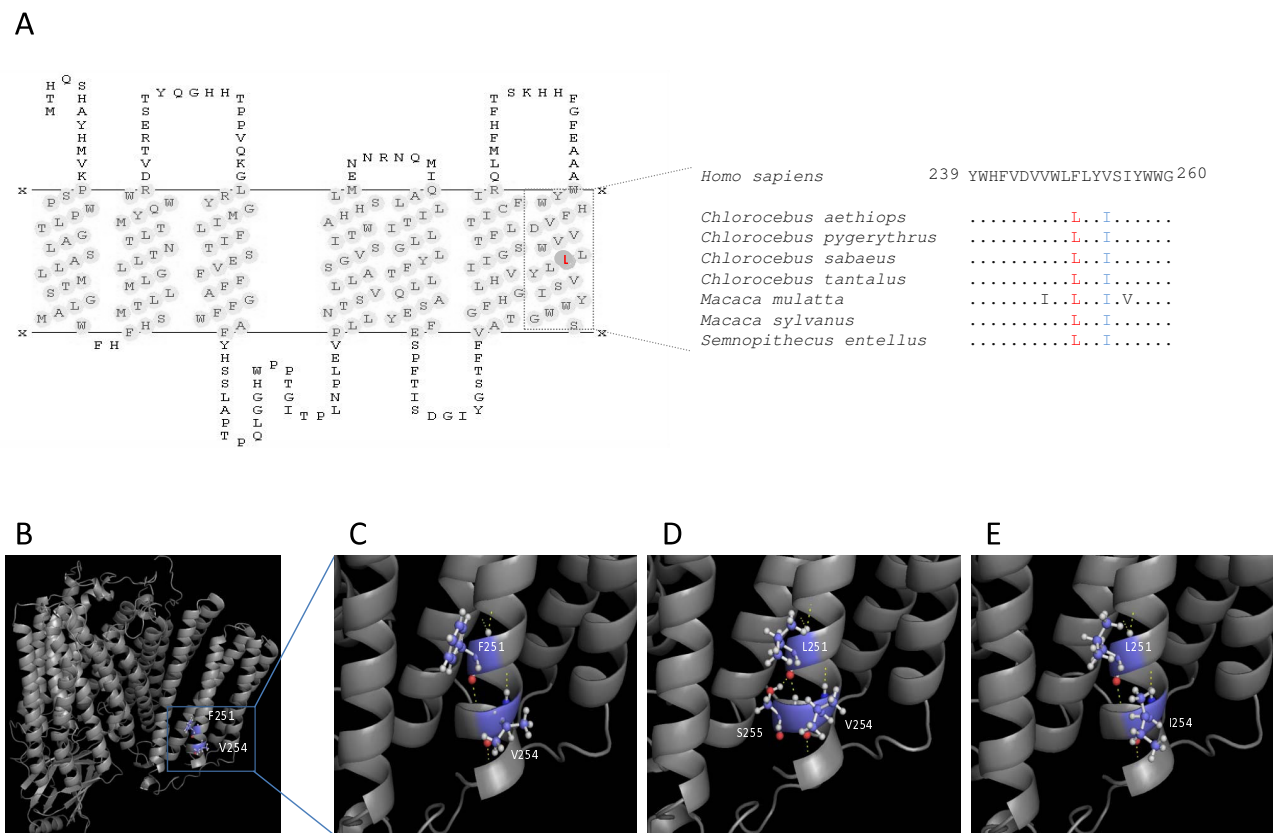
**Figure 1**  
**Molecular mechanism of compensation at COI.** (A) Three-dimensional model of human cytochrome oxidase c complex showing Leu195 and Leu196 positions in COI. (B) Detailed view of the wild-type background (195Leu-196Leu) showing side-chain bonding involving Ser187 ( $\alpha$ -helix V) and Leu248 ( $\alpha$ -helix VI). (C) Deleterious background (Leu195-Ile196) revealing the absence the H-bonding between  $\alpha$ -helices V and VI. (D) Compensated background (Ile195-Ile196) establishing a novel H-bond involving Ser198 ( $\alpha$ -helix V) and Phe238 ( $\alpha$ -helix VI). (E) Transmembrane structure of the human wild-type COI showing the mutation position (red), the interchain connection in the wild-type background (orange) and the new interaction in the compensated background (green).

Phe238) apart from the mutation (Ile196) and the compensatory residue itself (Ile195). Second, the compensatory residue does not intervene directly to maintain the connection between the two helices, although interaction with other residues seem to contribute to maintain proper protein folding [37]. The reestablishment of the original bonding pattern leading to the rescue of a deleterious phenotype is a remarkable observation in line with previous documented data [18].

For the identification of a compensatory partner for the COIII-Phe251Leu mutation we followed a similar approach. Comparative sequence analysis revealed 14 candidate compensatory residues for the residue 251 (Table 2), 13 of which are common to all of the species carrying the Leu251 mutation. Of these, position 254 emerged as the most likely compensatory site given the

spatial proximity in structural models (Figure 2A to 2C). These structures revealed that the human disease-associated background (Leu251-Val254, Figure 2D) holds an extra H-bond with Ser255 that is not seen in the wild-type background (Phe251-Val254, Figure 2C) and no longer observed in the compensated background (Leu251-Ile254, Figure 2E). Also, despite the amino-acid difference in positions 251 and 254 between wild-type and compensated background, the bonding pattern that involves both residues showed to be surprisingly similar (Figure 2C and Figure 2E).

Regarding the CYB-Gly251Ser mutation, interspecies sequence comparisons failed to reveal a compensatory residue that was shared by all species carrying the human deleterious Gly251 (Table 2) and this holds true even when placental mammals were considered separately

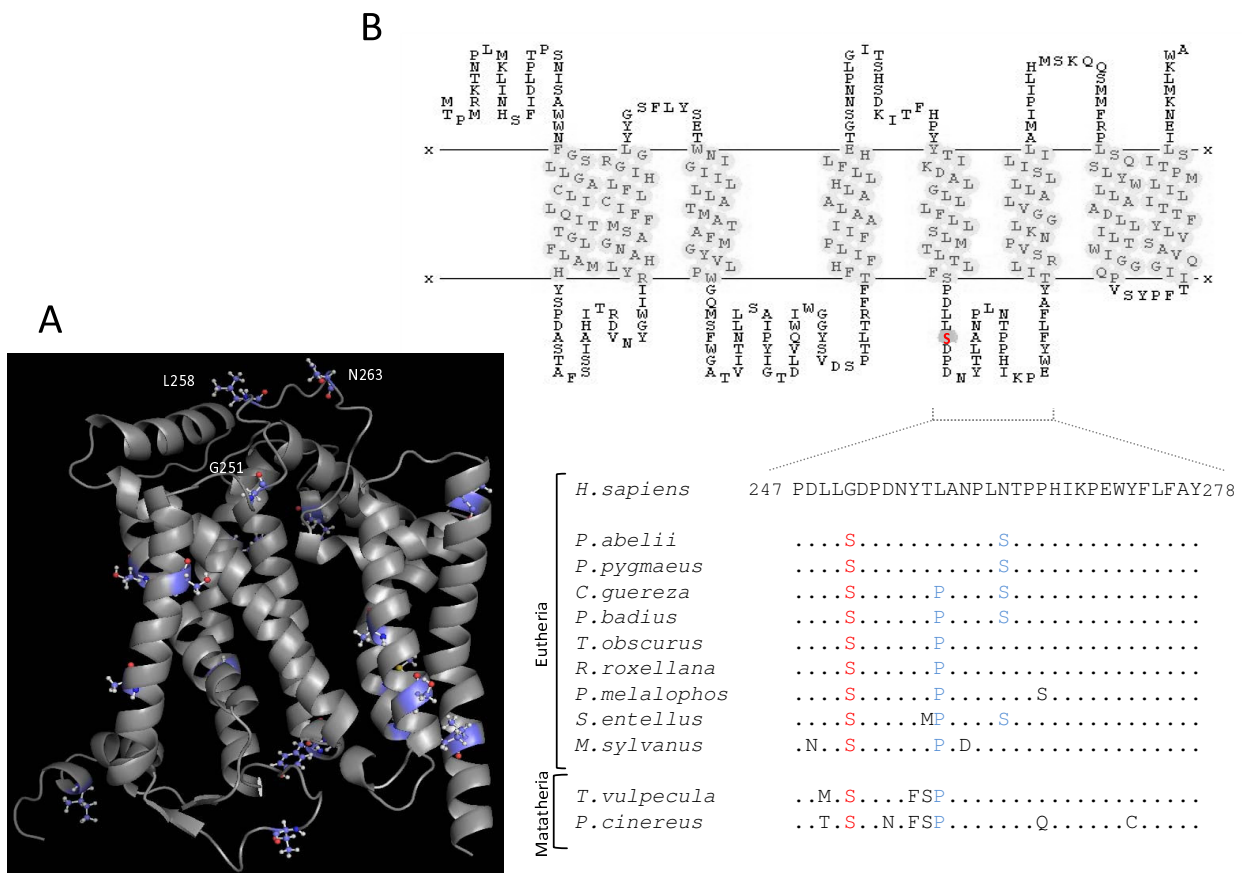


**Figure 2**  
**Molecular mechanism of compensation at COIII.** (A) Transmembrane structure highlighting the deleterious Leu251 at the C-terminus region and sequence comparison between human and primate species harboring the mutation and the most likely candidate to compensatory site. (B) Three-dimensional model of human cytochrome oxidase c complex showing positions Phe251 and Val254 at COIII. (C) H-bonding interactions between Phe251 and Val254 in the wild-type background. (D) Deleterious background (Leu251-Val254) showing a *de novo* interaction with Ser255. (E) Compensated background (Leu251-Ile254) illustrating the reestablishment of the original bonding pattern.

from marsupial lineages. Nevertheless, it is known that more than one compensatory site may exist for any given mutation [19] and thus phylogenetically related lineages can hold distinct compensatory solutions. We mapped onto the modeled 3D structure of CYB all of the possible compensatory alternatives (residues that differ between humans and all of the other mammals) although no structural proximity between each of these residues and Gly251 was evident (Figure 3A).

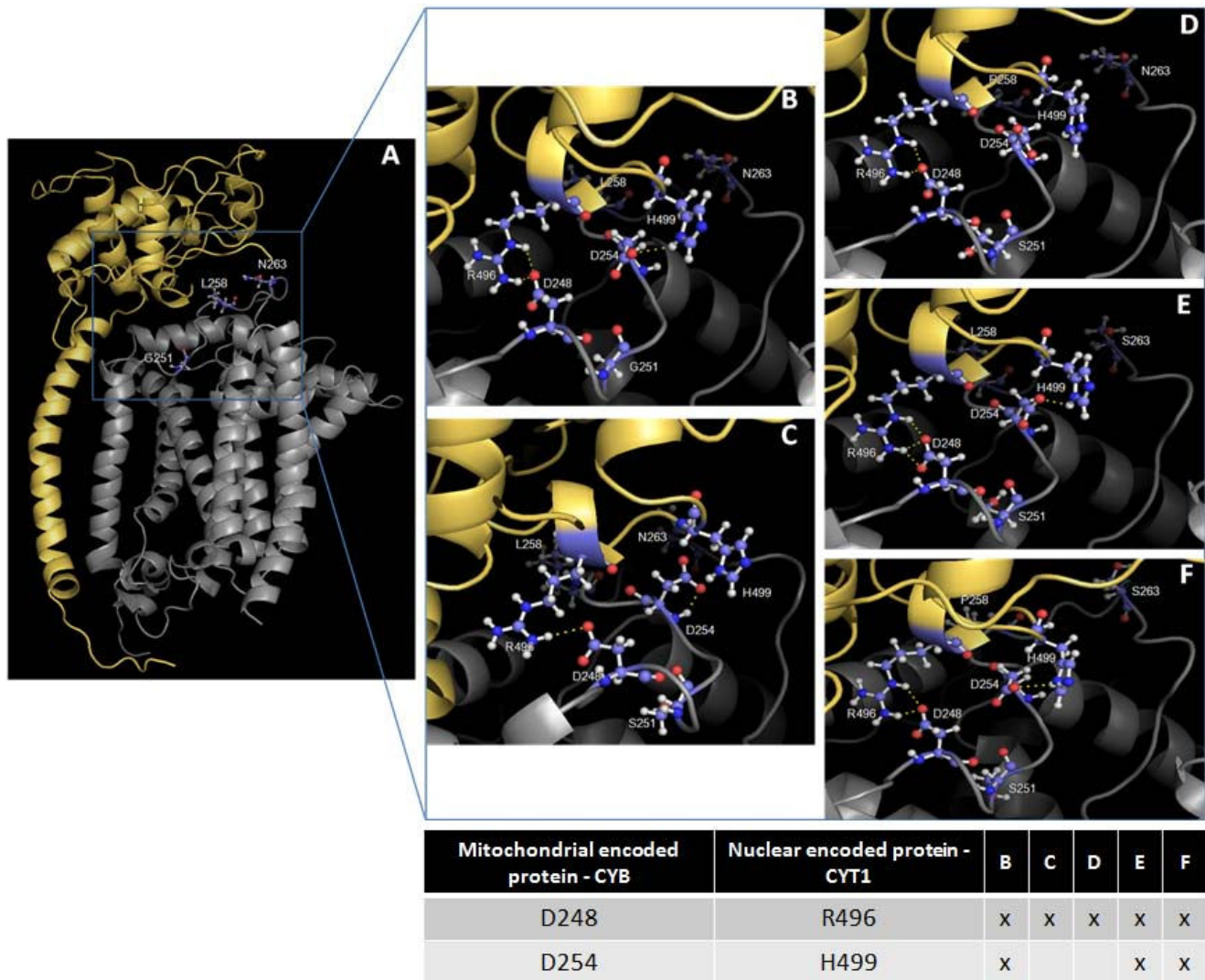
A previous study in yeast [18] revealed that mutations at the mitochondrial CYB can be reverted by distinct compensatory residues and also that each interacting pair lay on the same side of the membrane. To determine whether these evidence apply to the particular case of the Gly251Ser mutation, we focused on the loop that harbors the target residue 251 (Figure 3B). This domain revealed

two compensatory candidates, Pro258 and Ser263. Because no direct interaction between mutation and compensatory residue alone or in combination was evident (Figure 3A) and, as argued previously [38], mutations in this strongly conserved domain impairs the complex assembly, we hypothesized that the structural recovery of the mitochondrial *bc1* complex would require intermolecular interactions with other catalytic subunits, namely the nuclear-encoded cytochrome c1 (CYT1) [39]. In Figure 4, structures representing the mitochondrial-encoded CYB and the interacting region of CYT1 are shown for wild-type (Figure 4A and Figure 4B) and mutated human backgrounds (Figure 4C), as well as for all the possible combinations between the deleterious (251) and the compensatory 258 and 263 residues (Figure 4D to 4F).



**Figure 3 Molecular mechanism of compensation at CYB.** (A) Human CYB model highlighting the Gly251 residue and all putative compensatory sites [see Additional file 1]. (B) Transmembrane structure with the deleterious residue highlighted and comparison of human wild-type sequence with non-human carriers of the deleterious Ser251 (red) and compensatory Pro258 and Ser263 partners (blue).





**Figure 4**  
**Mitonuclear interaction between CYB and CYT1.** (A) Human model showing structural proximity between CYB (grey) and CYT1 (yellow). (B) Structure of the wild-type background illustrating H-bond salt bridges across CYB/CYT1 interfaces (CYB-Asp248/CYT1-Arg496 and CYB-Asp254/CYT1-His499). (C) Structure of the deleterious background (CYB-Ser251). (D), (E) and (F) Structure of the compensated backgrounds (CYB-Pro258, CYB-Ser263 and CYB-Pro258-Ser263, respectively). A synopsis of the CYB/CYT1 salt bridge bonding patterns for all of the backgrounds is provided below the structures.

Although protein-protein contact is unlikely to be conducted directly by the 258 and 263 residues, the region encompassing CYB/CYT1 interaction would involve H-bond salt bridges between the oppositely charged CYB-Asp248/CYT1-Arg496 and CYB-Asp254/CYT1-His499 residues (Figure 4B). These particular H-bonds are common across protein interfaces [40] and contribute to the stabilization of protein complexes [37,41]. In the mutated background (Figure 4C), one of these bonds (CYB-Asp254/CYT1-His499) seems to be missing, whereas it is recovered in two of the putative compensatory backgrounds harboring Ser263 alone or Ser263 in combina-

tion with Pro258 (Figure 4E and 4F). In the remaining background, which harbors Pro258 *solo*, this interaction was not restored (Figure 4D), similarly to what was observed in the deleterious background which may be explained by sequence differences in CYT1 between humans and the more distantly related mammals. These differences may contribute to structural recovery through interaction with the Pro258 residue. Unfortunately, no sequence data related to CYT1 is currently available for these non-human species to confirm this hypothesis. For that reason, this candidate was not included in the subsequent analyses.

To our knowledge, this study provides the first evidence for a compensatory interaction involving proteins encoded by mitochondrial and nuclear genomes in mammals, although a previous study focused on the fitness effect of mitonuclear epistatic interactions in *Drosophila melanogaster* [42] presented evidence that complement our results.

#### **Coupling of a deleterious mutation with a compensatory partner**

After the identification of four mutation-compensation pairs, we attempted to reconstruct the sequence of events that coupled a deleterious mutation with a compensatory partner.

The particular case of COI strongly suggested a quasi-instantaneous occurrence of the Ile196 mutation and compensatory Ile195, since either possible intermediate state (Leu195-Ile196 or Ile195-Leu196) could only have reached fixation at a prohibitively high fitness cost. In one case (Leu195-Ile196) the deleterious impact of the Leu196Ile substitution is well-documented [31]. The other possible intermediate state (Ile195-Leu196) would imply the replacement of an invariant residue (Leu195) in mammals, fishes, amphibians and reptiles, signaling its critical role to the protein function and put on evidence the constraints in replacing leucine for an alternative amino acid. Thus, each of these intermediate states would only be tolerated in heteroplasmy at low frequencies before reached fixation. It is worth noting that replaced residues (195 and 196) lie in the immediate vicinity but do not involve contiguous nucleotides. Thus, it is not obvious whether both nucleotide changes (the first position of each codon) have resulted from the same error during replication or, alternatively, have occurred sequentially towards the rapid fitness escape from unfavorable combinations [43]. But regardless of the mechanism of origin, a quasi-instantaneous fixation of the compensated background (homoplasmy) is possible to have occurred within a small number of generations [44,45].

The deleterious mutations in COIII and CYB were observed in multiple species allowing for the reconstruction of their evolutionary history using a perfectly resolved primate phylogenetic tree (Figure 5). We verified the presence of the candidate compensatory residues in several lineages without the deleterious mutations, but the opposite was never observed. Moreover, it was possible to track at least two independent events resulting in the deleterious COIII-Leu251, but again only in lineages harboring the corresponding compensatory solution. Our previous data also showed that a deleterious mutation in a nuclear-encoded protein (OTC) occurred independently in chimpanzees and dogs. In this case, the compensatory solution is the ancestral amino acid [10]. From these

examples, we are presenting a model of co-evolution in which genomes hold potential compensatory solutions for upcoming deleterious events.

In order to investigate whether other mammalian lineages could also have their genomes pre-compensating for COIII-Leu251 and CYB-Ser251 mutations the phylogenetic analysis was extended. The results, based on the analysis of all mammalian orders included in NCBI dataset, are graphically represented in Figure 6. The widespread occurrence of compensatory COIII-Ile254 and CYB-Ser263 across mammalian orders provides further support that these residues arose through independent mutational events. The remaining CYB-Pro258 residue is invariant in all mammalian orders, except for primates, and represents the ancestral residue as demonstrated by the perfectly resolved deepest branch of the tree (Figure 5B). This is similar to the case mentioned before [10].

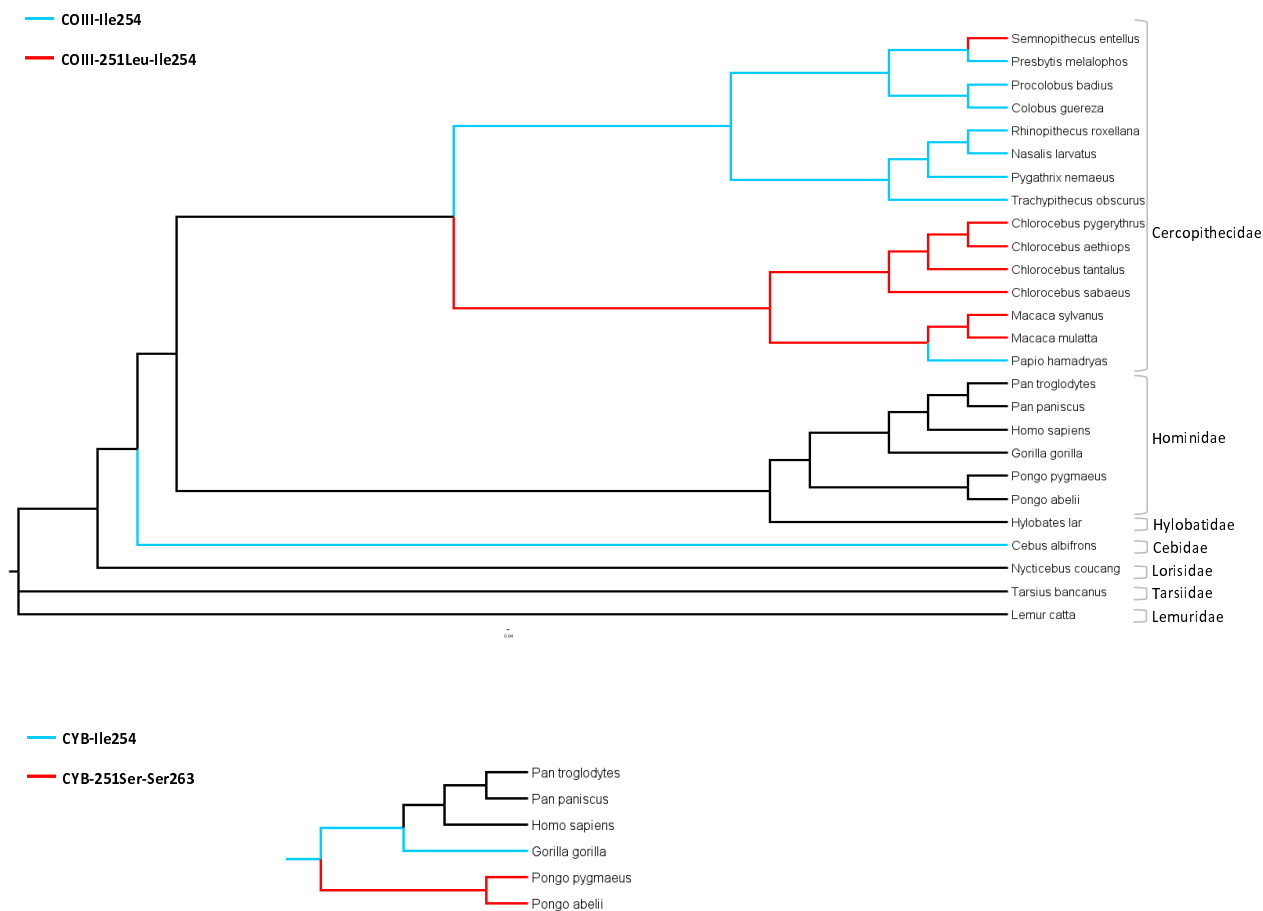
#### **Conclusion**

The role played by compensatory processes during protein evolution may now be understood at a greater level. This is aided by the emergence of more sophisticated theoretical and computational prediction tools and increasing empirical evidence. A mutation-compensation pair defines a structural and a functional example of coevolving sites in which the deleterious effect of the mutation is suppressed by its compensatory partner.

In this study, our efforts were directed towards the understanding of the molecular and evolutionary processes of compensation in three highly deleterious mutations that are the wild-type residue in non-human mammalian lineages. In two cases (COI-Leu196Ile and COIII-Phe251Leu), the reestablishment of the original bonding pattern in compensated backgrounds was observed through the extensive analysis of human modeled 3D structures. In the third case (CYB-Gly251Ser), no explicit interactions were observed between the mutation and any of the putative intramolecular compensatory partners. Nevertheless, as both are located in a critical region for the assembly of the *bc1* complex, we were able to infer the interaction between mitochondrial CYB and nuclear CYT1 in a noteworthy process of intermolecular epistatic interaction between the two genomes.

Although our analyses focused on mitochondrial-encoded proteins, the main conclusions may be extended to nuclear-encoded proteins as well. Under the assumption that a deleterious mutation occurs in pre-compensating nuclear background, one may explain how dominant and fitness-costly mutations are tolerated and reach fixation in various lineages. Furthermore, highly deleterious, yet recessive, mutations may persist in heterozygosity until a compensatory partner arises or becomes linked by





**Figure 5**  
**Maximum likelihood tree showing lineages harboring compensatory residues alone (blue) and compensated backgrounds (red).**

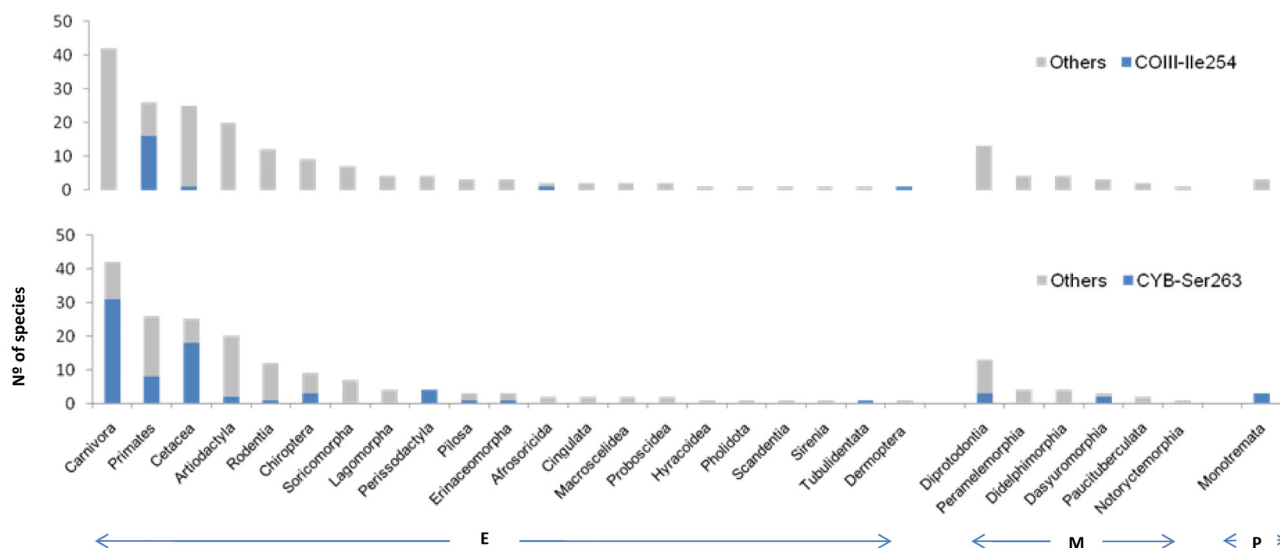
recombination. However, because even a small rate of recombination leads to increased fixation time of a compensated background under a strong selection pressure [25,26,46], the pre-compensating model would renders more likely the process of compensatory evolution, in particular when highly deleterious mutations are involved.

**Methods**

**Sequences and pathological data**

Mitochondrial-encoded protein sequences from over 230 mammalian species were extracted from NCBI Organelle Genome Resources [34]. Whenever data for more than one subspecies was available for each species, only one of those was considered resulting in a total of 199 distinctive species. Sequences were aligned with Clustal W software [47].

Mitochondrial mutations associated with non-LHON clinical phenotypes were obtained from the table "MtDNA Mutations with Reports of Disease-Associations-Coding & Control Region Mutations", available at Mitomap [48]. Only missense substitutions were considered. For all cases, the homologous site at the non-human protein was examined. This resulted in a total of nine cases where the human mutation overlaps the non-human homologous site in mammals (ND1-Met1Thr, ND1-Ala4Thr, ND1-Thr164Ala, ND2-Ala331Ser, ND4-Ile423Val, ND6-Val112Met, COI-Leu196Ile, COIII-Phe251Leu and CYB-Gly251Ser). Since no crystal structure was available for the respiratory complex I, only three cases, COI-L196I, COIII-F251L, and CYB-G251S were analyzed in the present study.



**Figure 6**  
**Distribution of pre-compensating mammalian genomes.** The Eutheria (E), Metatheria (M) and Prototheria (P) orders are represented in the X-axis. The number of species is presented at the Y-axis.

**Transmembrane structure model**

Transmembrane models were obtained from the Human Mitochondrial Genome Polymorphism Database [49].

**Protein comparative modeling and structure visualization**

We used the human sequence of COI (complex IV) and CYB (*bc1* complex) as templates to search for the highest E-value pdb using BlastP analysis [50]. The resulting codes of bovine structures (pdb 1occ and 1bgy for complex IV and *bc1*, respectively) were then used in MODELLER [51] in order to build human structural models. The accuracy of the predicted 3D human structural models [see Additional file 2] was estimated using ProSA-web [52], as previously described [53,54]. All of the structures here analyzed were built using MODELLER and are available at [http://www.portugene.com/data\\_Azevedo2008.html](http://www.portugene.com/data_Azevedo2008.html)[55]. All residue-residue bonds were calculated using Pymol software [56] and all of the structures were visualized using the same software.

**Phylogenetic inferences**

The primate phylogenetic tree was created using maximum likelihood calculations estimated with the PhyML algorithm [57] from an alignment of concatenated COIII and CYB protein sequences. The amino acid matrix used was mtRev and the substitution model was assumed to follow an approximate gamma distribution [58]. For testing the reliability of the PhyML tree topology, the original alignment file was converted to nexus format and served as input for the phylogenetic Bayesian software MrBayes

[59,60]. The amino acidic substitution model used was Mtmam following a gamma distribution [58,61,62]. Default settings were used for all of the remaining parameters. The statistical evaluation of posterior distribution for likelihood (LnL) was performed for the two independent MrBayes runs in Tracer v1.4 [63]. The resulting topology of the consensus Bayesian tree was identical to that generated by maximum likelihood analysis [see Additional file 3].

**Abbreviations**

mtDNA: mitochondrial DNA; NCBI: National Center for Biotechnology Information; OXPHOS: oxidative phosphorylation; COI: cytochrome c oxidase I; COIII: cytochrome c oxidase III; CYB: cytochrome b; CYT1: cytochrome c1

**Authors' contributions**

LA, JC and AA conceived the study and main analyses. LA, JC and AM carried out all the analysis. LA, JC, BA, AM, FP and AA analyzed and interpreted the data. LA, JC, BA and AA wrote the manuscript. All authors read and approved the final manuscript.

## Additional material

### Additional file 1

*Candidate compensatory residues for human deleterious mutations. This table provides all the possible compensatory sites for three human deleterious mutations found in non-human mammals.*

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-266-S1.doc>]

### Additional file 2

*Quality evaluation of the modeled structures of human mitochondrial proteins. This figure presents the evaluation of model quality for the predicted 3D structures of human COI/COIII (A), CYB (B) and CYT1 (C) both overall (left) and locally (right) as estimated in ProSA-web. Both modeled structures showed z score values (black dots) that lie within the cloud, representing experimentally determined features of native proteins by X-ray and NMR analysis. Energy plots show a smooth fluctuation with overall negative energy of residue stretches (green lines) demonstrating that the predicted 3D structures show minimal deviations from normal energy values.*

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-266-S2.tiff>]

### Additional file 3

*Consensus phylogeny of primate lineages. The figure shows the MrBayes consensus tree illustrating primate topology (A), the marginal density of posterior distribution of likelihood (LnL) for first and second MrBayes runs (B) and Tracer statistical results for tree likelihood, TL (tree length) and alpha in first and second run of MrBayes (C).*

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-266-S3.tiff>]

## Acknowledgements

The authors acknowledge the comments of two anonymous reviewers. This work was supported by the research project PTDC/CVT/64154/2006. LA is supported by Fundação para a Ciência e a Tecnologia (FCT) Ciência 2007. BA (SFRH/BDE/15581/2006) and FP (SFRH/BD/19585/2004) are grantees from the FCT. IPATIMUP is partially supported by "Programa Operacional Ciência e Inovação 2010" (POCI 2010), VI Programa Quadro (2002–2006).

## References

- Kimura M: **The role of compensatory neutral mutations in molecular evolution.** *J Genet* 1985, **7**:19.
- Kern AD, Kondrashov FA: **Mechanisms and convergence of compensatory evolution in mammalian mitochondrial tRNAs.** *Nat Genet* 2004, **36**:1207-1212.
- Parsch J, Tanda S, Stephan W: **Site-directed mutations reveal long-range compensatory interactions in the Adh gene of Drosophila melanogaster.** *Proc Natl Acad Sci USA* 1997, **94**:928-933.
- Savill NJ, Hoyle DC, Higgs PG: **RNA sequence evolution with secondary structure constraints: comparison of substitution rate models using maximum-likelihood methods.** *Genetics* 2001, **157**:399-411.
- Stephan W, Kirby DA: **RNA folding in Drosophila shows a distance effect for compensatory fitness interactions.** *Genetics* 1993, **135**:97-103.
- Wilke CO, Lenski RE, Adami C: **Compensatory mutations cause excess of antagonistic epistasis in RNA secondary structure folding.** *BMC Evol Biol* 2003, **3**:3.
- Chen Y, Stephan W: **Compensatory evolution of a precursor messenger RNA secondary structure in the Drosophila melanogaster Adh gene.** *Proc Natl Acad Sci USA* 2003, **100**:11499-11504.
- Howe KJ, Ares M: **Intron self-complementarity enforces exon inclusion in a yeast pre-mRNA.** *Proc Natl Acad Sci USA* 1997, **94**:12467-12472.
- Poon A, Chao L: **Functional origins of the fitness effect-sizes of compensatory mutations in the DNA bacteriophage {phi}X174.** *Evolution* 2006, **60**:2032-2043.
- Azevedo L, Suriano G, van Asch B, Harding RM, Amorim A: **Epistatic interactions: how strong in disease and evolution?** *Trends Genet* 2006, **22**:581-585.
- Conner JR, Smirnova II, Peltorak A: **Forward genetic analysis of Toll-like receptor responses in wild-derived mice reveals a novel antiinflammatory role for IRAK1BP1.** *J Exp Med* 2008, **205**:305-314.
- Kondrashov AS, Sunyaev S, Kondrashov FA: **Dobzhansky-Muller incompatibilities in protein evolution.** *Proc Natl Acad Sci USA* 2002, **99**:14878-14883.
- Kulathinal RJ, Bettencourt BR, Hartl DL: **Compensated deleterious mutations in insect genomes.** *Science* 2004, **306**:1553-1554.
- Suriano G, Azevedo L, Novais M, Boscolo B, Seruca R, Amorim A, Ghibaudi EM: **In vitro demonstration of intra-locus compensation using the ornithine transcarbamylase protein as model.** *Hum Mol Genet* 2007, **16**:2209-2214.
- Mateu MG, Fersht AR: **Mutually compensatory mutations during evolution of the tetramerization domain of tumor suppressor p53 lead to impaired hetero-oligomerization.** *Proc Natl Acad Sci USA* 1999, **96**:3595-3599.
- Wilson KP, Malcolm BA, Matthews BW: **Structural and thermodynamic analysis of compensating mutations within the core of chicken egg white lysozyme.** *J Biol Chem* 1992, **267**:10842-10849.
- Moore JH: **A global view of epistasis.** *Nat Genet* 2005, **37**:13-14.
- di Rago J-P, Hermann-Le Denmat S, Pâques F, Rislisler J-I, Netter P, Słonimski PP: **Genetic Analysis of the Folded Structure of Yeast Mitochondrial Cytochrome b Selection of Intragenic Second-site Revertants.** *Journal of Molecular Biology* 1995, **248**:804-811.
- Poon A, Davis BH, Chao L: **The coupon collector and the suppressor mutation: estimating the number of compensatory mutations by maximum likelihood.** *Genetics* 2005, **170**:1323-1332.
- Poon A, Chao L: **The Rate of Compensatory Mutation in the DNA Bacteriophage {phi}X174.** *Genetics* 2005, **170**:989-999.
- Ferrer-Costa C, Orozco M, Cruz X: **Characterization of Compensated Mutations in Terms of Structural and Physico-Chemical Properties.** *Journal of Molecular Biology* 2007, **365**:249-256.
- Subramanian S, Kumar S: **Evolutionary anatomies of positions and types of disease-associated and neutral amino acid mutations in the human genome.** *BMC Genomics* 2006, **7**:306.
- DePristo MA, Weinreich DM, Hartl DL: **Missense meanderings in sequence space: a biophysical view of protein evolution.** *Nat Rev Genet* 2005, **6**:678-687.
- Gao L, Zhang J: **Why are some human disease-associated mutations fixed in mice?** *Trends in Genetics* 2003, **19**:678-681.
- Innan H, Stephan W: **Selection intensity against deleterious mutations in RNA secondary structures and rate of compensatory nucleotide substitutions.** *Genetics* 2001, **159**:389-399.
- Stephan W: **The Rate of Compensatory Evolution.** *Genetics* 1996, **144**:419-426.
- Hartl DL, Taubes CH: **Compensatory Nearly Neutral Mutations: Selection without Adaptation.** *Journal of Theoretical Biology* 1996, **182**:303-309.
- Rand DM, Haney RA, Fry AJ: **Cytonuclear coevolution: the genomics of cooperation.** *Trends in Ecology & Evolution* 2004, **19**:645-653.
- Manfredi G, Schon EA, Moraes CT, Bonilla E, Berry GT, Sladky JT, DiMauro S: **A new mutation associated with MELAS is located in a mitochondrial DNA polypeptide-coding gene.** *Neuromuscul Disord* 1995, **5**:391-398.

30. Tarnopolsky MA, Simon DK, Roy BD, Chorneyko K, Lowther SA, Johns DR, Sandhu JK, Li Y, Sikorska M: **Attenuation of free radical production and paracrystalline inclusions by creatine supplementation in a patient with a novel cytochrome b mutation.** *Muscle Nerve* 2004:537-547.
31. Varlamov DA, Kudin AP, Vielhaber S, Schrader R, Sassen R, Becker A, Kunz D, Haug K, Rebstock J, Heils A, et al.: **Metabolic consequences of a novel missense mutation of the mtDNA CO I gene.** *Hum Mol Genet* 2002, 11:1797-1805.
32. Klein G, Georgopoulos C: **Identification of Important Amino Acid Residues That Modulate Binding of Escherichia coli GroEL to Its Various Cochaperones.** *Genetics* 2001, 158:507-517.
33. Martin LC, Gloor GB, Dunn SD, Wahl LM: **Using information theory to search for co-evolving residues in proteins.** *Bioinformatics* 2005, 21:4116-4124.
34. Li Y, Wray R, Blount P: **Intragenic suppression of gain-of-function mutations in the Escherichia coli mechanosensitive channel, MscL.** *Molecular Microbiology* 2004, 53:485-495.
35. Iwata S, Lee JW, Okada K, Lee JK, Iwata M, Rasmussen B, Link TA, Ramaswamy S, Jap BK: **Complete Structure of the 11-Subunit Bovine Mitochondrial Cytochrome bc1 Complex.** *Science* 1998, 281:64-71.
36. Tsukihara T, Aoyama H, Yamashita E, Tomizaki T, Yamaguchi H, Shinzawa-Ittoh K, Nakashima R, Yaono R, Yoshikawa S: **The Whole Structure of the 13-Subunit Oxidized Cytochrome c Oxidase at 2.8 Å.** *Science* 1996, 272:1136-1144.
37. Xu D, Lin SL, Nussinov R: **Protein binding versus protein folding: the role of hydrophilic bridges in protein associations.** *Journal of Molecular Biology* 1997, 265:68-84.
38. Brasseur G, Sariba AS, Daldal F: **A compilation of mutations located in the cytochrome b subunit of the bacterial and mitochondrial bc1 complex.** *Biochim Biophys Acta* 1996, 1275:61-69.
39. Saint-Georges Y, Bonnefoy N, di Rago JP, Chiron S, Dujardin G: **A Pathogenic Cytochrome b Mutation Reveals New Interactions between Subunits of the Mitochondrial bc1 Complex.** *J Biol Chem* 2002, 277:49397-49402.
40. Xu D, Tsai CJ, Nussinov R: **Hydrogen bonds and salt bridges across protein-protein interfaces.** *Protein Eng* 1997, 10:999-1012.
41. Sheinerman FB, Norel R, Honig B: **Electrostatic aspects of protein-protein interactions.** *Current Opinion in Structural Biology* 2000, 10:153-159.
42. Dowling DK, Friberg U, Hailer F, Arnqvist G: **Intergenomic Epistasis for Fitness: Within-Population Interactions Between Cytoplasmic and Nuclear Genes in Drosophila melanogaster.** *Genetics* 2007, 175:235-244.
43. Weinreich DM, Chao L: **Rapid evolutionary escape by large populations from local fitness peaks is likely in nature.** *Evolution* 2005, 59:1175-1182.
44. Ashley MV, Laipis PJ, Hauswirth WW: **Rapid segregation of heteroplasmic bovine mitochondria.** *Nucleic Acids Res* 1989, 17:7325-7331.
45. Hauswirth WW, Laipis PJ: **Mitochondrial DNA polymorphism in a maternal lineage of Holstein cows.** *Proc Natl Acad Sci USA* 1982, 79:4686-4690.
46. Higgs P: **Compensatory neutral mutations and the evolution of RNA.** *Genetica* 1998, 102-103:91-101.
47. **Clustal W** [<http://www.ebi.ac.uk/tools/clustalw2>]
48. **MITOMAP: A Human Mitochondrial Genome Database** [<http://www.mitomap.org>]
49. **mtSNP: Human Mitochondrial Genome Polymorphism Database** [[http://mtsnip.tmg.or.jp/mtsnip/index\\_e.shtml](http://mtsnip.tmg.or.jp/mtsnip/index_e.shtml)]
50. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, 25:3389-3402.
51. Sali A, Blundell TL: **Comparative protein modeling by satisfaction of spatial restraints.** *J Mol Biol* 1993, 234:779-815.
52. **ProSA-web - Protein Structure Analyses** [<https://prosa.services.came.sbg.ac.at/prosa.php>]
53. Sippl MJ: **Recognition of errors in three-dimensional structures of proteins.** *Proteins* 1993, 17:355-362.
54. Wiederstein M, Sippl MJ: **ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins.** *Nucleic Acids Res* 2007, 35:W407-410.
55. **Modeled structures** [<http://www.portugene.com/data/Azevedo2008.html>]
56. **Pymol software** [<http://www.pymol.org>]
57. Guindon S, Gascuel O: **A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood.** *Syst Biol* 2003, 52:696-704.
58. Yang Z, Nielsen R, Hasegawa M: **Models of amino acid substitution and applications to mitochondrial protein evolution.** *Mol Biol Evol* 1998, 15:1600-1611.
59. Huelsenbeck JP, Ronquist F: **MRBAYES: Bayesian inference of phylogenetic trees.** *Bioinformatics* 2001, 17:754-755.
60. Ronquist F, Huelsenbeck JP: **MrBayes 3: Bayesian phylogenetic inference under mixed models.** *Bioinformatics* 2003, 19:1572-1574.
61. Cao Y, Janke A, Waddell PJ, Westerman M, Takenaka O, Murata S, Okada N, Pääbo S, Hasegawa M: **Conflict Among Individual Mitochondrial Proteins in Resolving the Phylogeny of Eutherian Orders.** *Journal of Molecular Evolution* 1998, 47:307-322.
62. Yang Z: **Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites.** *Mol Biol Evol* 1993, 10:1396-1401.
63. **Tracer v1.4** [<http://beast.bio.ed.ac.uk/Tracer>]

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

