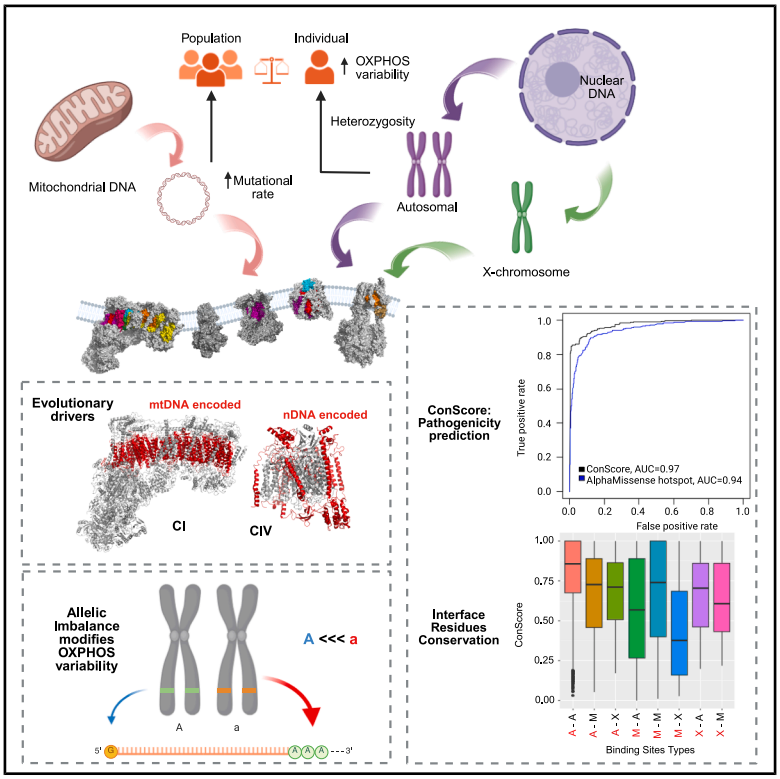
Cell Genomics

Article

Structural diversity and evolutionary constraints of oxidative phosphorylation

Graphical abstract



Highlights

- Respiratory complex evolution is driven by mtDNA for CI or by nuclear genes for CIV
- Interface residues in mtDNA-encoded subunits show unexpectedly lower conservation
- Allelic imbalance emerges as a novel mechanism limiting OxPhos genetic variability
- ConScore integrates multi-scale conservation data to predict OxPhos pathogenic variants

Authors

José Luis Cabrera-Alarcón, Marina Rosa-Moreno, Lucía Sánchez-García, ..., Fernando Martínez, Fátima Sánchez-Cabo, José Antonio Enríquez

Correspondence

jaenriquez@cnic.es

In brief

Cabrera-Alarcón et al. reveal that, contrary to prior assumptions, the oxidative phosphorylation system exhibits significant genetic variability within individuals and populations. By integrating evolutionary, structural, and genetic analyses, they uncover mechanisms that constrain variability and introduce a predictive tool, offering new perspectives on OxPhos function, speciation, and disease.

Cabrera-Alarcón et al., 2025, Cell Genomics 5, 100945
September 10, 2025
https://doi.org/10.1016/j.xgen.2025.100945

CellPress

Cell Genomics

CellPress
OPEN ACCESS

Article

Structural diversity and evolutionary constraints of oxidative phosphorylation

José Luis Cabrera-Alarcón,[1,2] Marina Rosa-Moreno,[1] Lucía Sánchez-García,[3] Pablo Hernansanz-Agustín,[1,2] Maria Concepción Jiménez-Gómez,[1,2] Fernando Martínez,[4] Fátima Sánchez-Cabo,[3,4,5] and José Antonio Enríquez[1,2,6,*]

[1]GENOXPHOS Lab, Centro Nacional de Investigaciones Cardiovasculares (CNIC), 28029 Madrid, Spain
[2]Centro de Investigaciones Biomédicas en Red en Fragilidad y Envejecimiento Saludable (CIBERFES), 28029 Madrid, Spain
[3]Computational Systems Biomedicine Lab, Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain
[4]Bioinformatics Unit, Computational Systems Biomedicine Lab, Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain
[5]Centro de Investigaciones Biomédicas en Red en Investigación Cardiovascular (CIBER-CV), 28029 Madrid, Spain
[6]Lead contact
*Correspondence: jaenriquez@cnic.es
https://doi.org/10.1016/j.xgen.2025.100945

SUMMARY

The oxidative phosphorylation (OxPhos) system is central to metabolism. The more than 90 structural subunits are encoded by different chromosome categories (autosomal, X, and mtDNA). The system is envisioned as an invariant structure between cells and individuals. However, a comprehensive analysis of the 1,000 Genomes Project data reveals unexpected genetic intra-individual variability resulting from the heterozygosity of diploid autosomal genes, while diversity at the population level is generated by variability in mtDNA. We characterized the different levels of structural constriction at evolutionary and population levels for all OxPhos protein residues. To support this analysis, we developed ConScore, a conservation-based predictor of variant impact within OxPhos proteins (area under the receiver operating characteristic curve [ROC-AUC] = 0.97; area under the precision-recall curve [PR-AUC] = 0.94). Notably, for the nuclear-encoded subunits, we found mechanisms limiting individual variability as allelic imbalance or homozygosity bias. Integrating structural, functional, and genetic data, we highlight the significance of each OxPhos protein position, expanding insights into its role in speciation and disease.

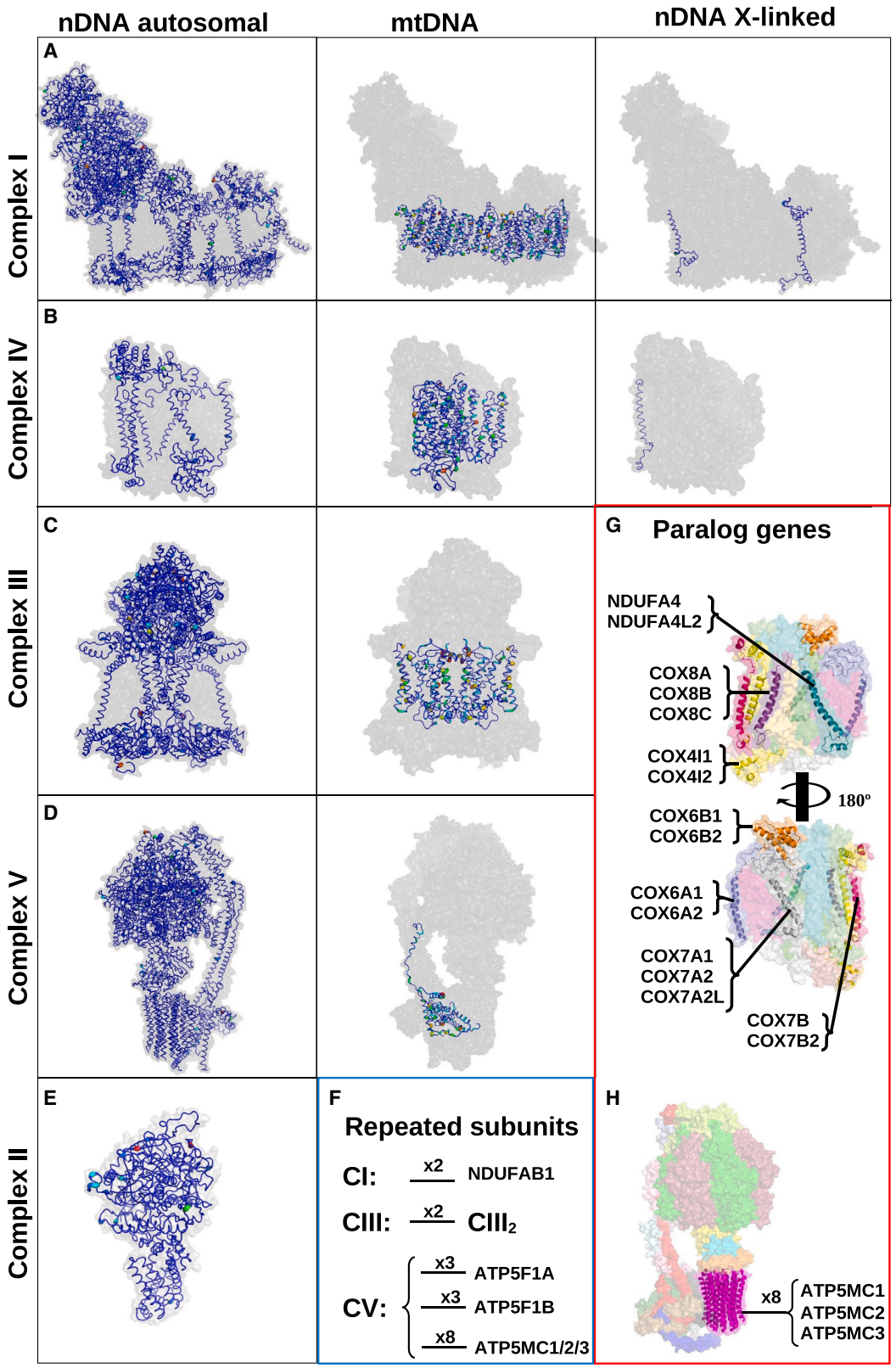
INTRODUCTION

Respiratory complexes (RCs) comprise subunits originating from mitochondrial DNA (mtDNA) and nuclear DNA (nDNA), with 13 mtDNA-encoded and 77 nDNA-encoded subunits across the five complexes (CI–CV).[1] Three nDNA subunits (Ndufa1, Cox7b, and Ndufb11) derive from X chromosome genes (Figures 1A–1E). Structural complexity arises from subunit repetition (e.g., two NDUFAB1 in CI, dimeric CIII, multiple ATP5F1A/B copies in CV) (Figure 1F) and paralogous isoform incorporation in CIV and CV (Figures 1G and 1H), expanding total components to 103.

Functional assembly requires precise subunit compatibility. Studies show mtDNA haplotype variations paired with identical nDNA alter RC composition, supercomplex formation, and metabolic output, indicating adaptable OxPhos performance within healthy ranges.[2] However, coexisting mtDNA haplotypes disrupt OxPhos efficiency, suggesting excessive heterogeneity is detrimental.[3] Such mitonuclear incompatibilities, particularly in allopatric populations, contribute to speciation by impairing hybrid viability. Incompatibilities operate at multiple levels, including pre-copulatory mechanisms.[4] Mitonuclear incompatibilities are not limited to OxPhos system components.[5] However, the OxPhos system remains the most direct manifestation of mitonuclear incompatibility. This concept is the foundational basis for our work, analyzing at the level of residue the evolutionary strategies within the OxPhos system subunits and how genetic diversity is regulated in the diploid nuclear-encoded OxPhos genes. To address this question, we analyzed the scope of OxPhos variability in the individuals included in the 1000 Genomes Project.[6] We specifically quantify the significance of heterozygosity on individual structural variability. Our investigation unearths significant interindividual diversity in normal human populations based on genotype arising from the biallelic autosomal-encoded subunits. We also found that variability arising from diploidy might also be controlled by mechanisms regulating the differential expression of heterozygote alleles.

We explore how genetic evolution and variability in the OxPhos system are integrated in a functional and structural context to understand the adaptive forces that have shaped mitonuclear evolutionary dynamics and also how evolutionary strategies have diversified from vertebrates to current human populations according to the different RCs or interaction between autosomal, mtDNA, or X-linked genes, recognizing that evolution can act with fine granularity at the level of individual residues.

To address such questions, this study examines population variability in OxPhos genes from the 1000 Genomes Project[6] and gnomAD-v.3,[7] together with multiple sequence alignments



(legend on next page)

(MSAs) from vertebrate, mammal, and primate species. The analysis focused on the potential for the assembly of alternative versions of RCs, the degree of mutability at individual protein positions, and the implications of these findings for proteins encoded by distinct chromosome types. We also explored the role of allelic imbalance in shaping the OxPhos landscape. In RCs, as in any other protein, function is determined by structure. Therefore, the importance of variability at any specific position relies on the three-dimensional (3D) structure. Thus, our study defines a 3D map of the relative importance of all individual amino acid residues in RCs based on evolutionary and population information embedded in RC structures. The integration of population genetics, structural biology, and evolutionary analysis provides a comprehensive view of how this essential metabolic system has evolved and adapted across different species while maintaining functional integrity.

RESULTS

Structural development and high-throughput refinement of human OxPhos RCs

The 3D protein structures allow mapping the location of mutations and assessing their impact on the function of the OxPhos system, considering aspects such as structural resilience and functional compatibility. Thus, mapping the subsequent analyses onto reliable structures will enable the creation of a four-dimensional framework (with evolution as the fourth dimension), allowing for multi-level analysis from residue to RC, including protein positions, subunits, subunit types, and the entire RC. This approach requires well-curated structural models, the creation of new models, and the refinement of existing ones. For CII and CV, we had to develop models based on existing structures from other species using the Rosetta *in silico* modeling tool suite.[8–10] For CI, CIII, and CIV, the published structures for humans needed to be substantially refined to resolve missing fragments and to eliminate steric clashes. We achieved this refinement through energy minimization, which in several instances resulted in a fine-tuning of the backbone positioning.

Genotype data retrieved from a large database reveal great heterogeneity in the landscape of RCs at the individual level

To determine how many different OxPhos complexes an individual can assemble, we analyzed the extent of individual genetic diversity of the OxPhos system subunits in all the 2,504 individualized genotypes included in the 1000 Genomes Project, considering the alleles that differ by one or more amino acids (Figure 2). The variability due to diploidy + heterozygosity correlates with the number of subunits in each complex. Thus, more than 90% of individuals had the potential to assemble at least two versions of CI, and, while most had only one possible version of CII (70%) and CIII (53%), a significant number of individuals could potentially form multiple subunit combinations for these RCs. For CIV and CV, the existence of paralog genes encoding alternative subunit isoforms increases the number of possible configurations, and no individual was able to form only one, underscoring the way in which alternative isoforms enlarge the base pool of possible versions of these RCs (Figure 2).

In CIV, the potential for additional variability due to diploidy is marginal compared with the other complexes, indicating that a unique adaptive strategy holds for the oxygen consuming RC. Nevertheless, heterozygosity of CIV subunits increased the number of possible CIV versions in about 33% of individuals.

Our analysis underscores the complexity and adaptability of human biochemistry at the molecular level and reveals that heterozygosity in nuclear OxPhos subunits serves as a significant source of intra-individual variability. However, for CIV, while this factor is present, most variability arises from the utilization of isoforms encoded by paralogous genes.

Evolutionary strategies for each OxPhos complex

We next inquire which residues are free to vary and whether this reflects a specific evolutionary trajectory that has shaped OxPhos system. Thus, we investigated the evolutionary trajectory of OxPhos components' positions to find evolutionary dynamics conditioned by the class of the coding chromosome (mtDNA, X, or autosomal) and to identify which RCs are more involved in the speciation process. The residue conservation rate of OxPhos proteins was determined in nested phylogenetic groups, vertebrates, mammals, and primates using MSA. Then, from these MSAs, a conservation score at each position (each column in the MSA) was determined by Shannon entropy position variation (SEPV). In addition, residue conservation among human populations (conservation at intraspecific level, as current step of evolution process) was also calculated as SEPV, using the aggregated genome information from more than 76,000 individuals in gnomAD-v3 frequency database. Next, SEPV values were minimum-maximum (min-max) normalized in each evolutionary point to transform data to the same scale. In this common scale, the closer SEPV to 0 the more conserved, while the closer to 1 the more variable. In consequence, SEPV analysis at these four groups will allow us to uncover patterns linked to purifying selection, local adaptation, and genetic equilibrium. This analysis was performed for all OxPhos subunits, and an SEPV conservation score was calculated for each position in each RC, considering the stoichiometric composition in each case.

Conservation study by RC

First, we explored the conservation differences between RCs across these four phylogenetic groups (Figure 3A). CV and CII were the most conserved RCs in all groups. CV has the highest number of repeated subunits (Figure 1), increasing their selective pressure, since incompatibilities are difficult to compensate by neighbors. CII is likely conserved due to its small size and the lack of mtDNA-encoded subunits. CI and CIV are the most variable between species. However, their evolution varies in the different groups. CIII is the main source of variation between

Figure 1. The structural complexity of OxPhos multimeric complexes

(A–E) Structural representation of OxPhos proteins according to coding chromosomes (autosomal, mtDNA, or X).

(F) Repeated subunits within the minimum complexity conformations of CI, CIII, and CIV.

(G and H) Positions occupied by isoforms encoded by paralog genes in CIV (G) and CV (H).

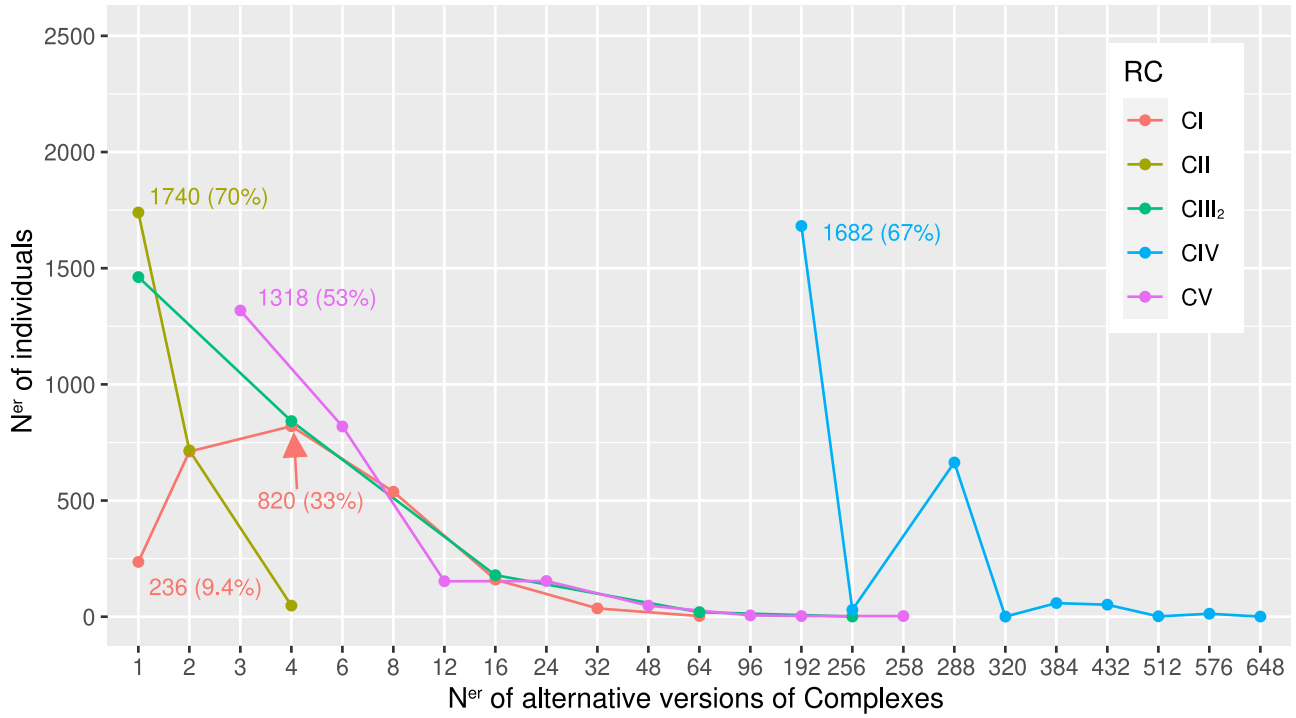


Figure 2. Number of potentially different RCs assembled per individual genotyped in the 1,000 Genomes Project (*n* = 2,504)

Each point represents the number of individuals (*y* axis) in relation to the number of different complexes that can potentially be assembled, given their genotypes (*x* axis).

vertebrates, while CI and CIV become the main evolutionary drivers from mammals onward. Notice that CI, CIII, and CIV are responsible for pumping protons across the inner mitochondrial membrane, a function in which the subunits encoded by the mtDNA play a key role. Moreover, these three RCs are involved in the formation of supercomplexes.

Conservation assessment by subunit inheritance type

Considering the whole OxPhos system in vertebrates, mammals, and primates, the mtDNA and X-linked subunits are the main source of variation. Conversely the X-linked subunits are the most conserved in humans, aligning with the Haldane effect (Figure 3B). Considering the accumulation of incompatibilities a process that drives speciation, X-linked genes become very relevant because of hemizygosity and monoallelic expression. Interestingly, within vertebrates and in some mammals, human X-encoded OxPhos genes lose their X-linked condition. Moreover, X-linked subunits are less conserved than mtDNA-encoded subunits in vertebrates, mammals, and primates. In addition, mtDNA-encoded subunits represent 4,169 positions versus 303 for X-linked subunits. Altogether, with the higher mutation rate of mtDNA, this maintains its role as a source of variability in humans as the main evolutionary drivers of the OxPhos system.

The RCs constitute independent entities despite being part of the same system. Each one will have its own particularities, repetition of subunits in its elementary structure, dimeric presentation as basic structural form, alternative isoforms encoded by different genes, have or not have subunits encoded by the mtDNA, representing different proportions within the RC, or contain X-linked subunits in its structure. Therefore, each RC is unique and may have adopted a particular evolutionary strategy. Therefore, we next analyze the conservation of the different OxPhos subunit types (autosomal, X-linked, and mtDNA-encoded subunits) by RC.

Conservation based on subunit inheritance type in CI

CI is one of the main electron inputs to the ETC, as well as one of the main players involved in both physiological and pathological production of reactive oxygen species (ROS), and is directly involved in the maintenance of membrane potential through H^+ pumping and as Na^+/H^+ antiporter.[11,12] In addition, it is the RC that contains the highest number of mtDNA-encoded subunits (7 MT-ND1-6 and MT-ND4L) and the highest number of X-linked subunits (NDUFA1 and NDUFB11). Consequently, both mtDNA-encoded and X-linked subunits constitute the main sources of evolutionary change in vertebrates, mammals, and primates (Figure 3C). However, as previously mentioned, the X-linked subunits among human populations are more conserved in response to the Haldane effect, leaving the seven mtDNA-encoded subunits as the main source of variation between individuals for CI (Figure 3C right panel).

Conservation by subunit inheritance type in CIII

CIII, the other major player in ROS production, is naturally found as a constituent of the OxPhos system forming dimers.[1] MT-CYB is the only CIII subunit encoded by the mtDNA, which is involved in both the Q-cycle and coupled proton pumping.[13] In CIII, MT-CYB represents the main source of change at all evolutionary

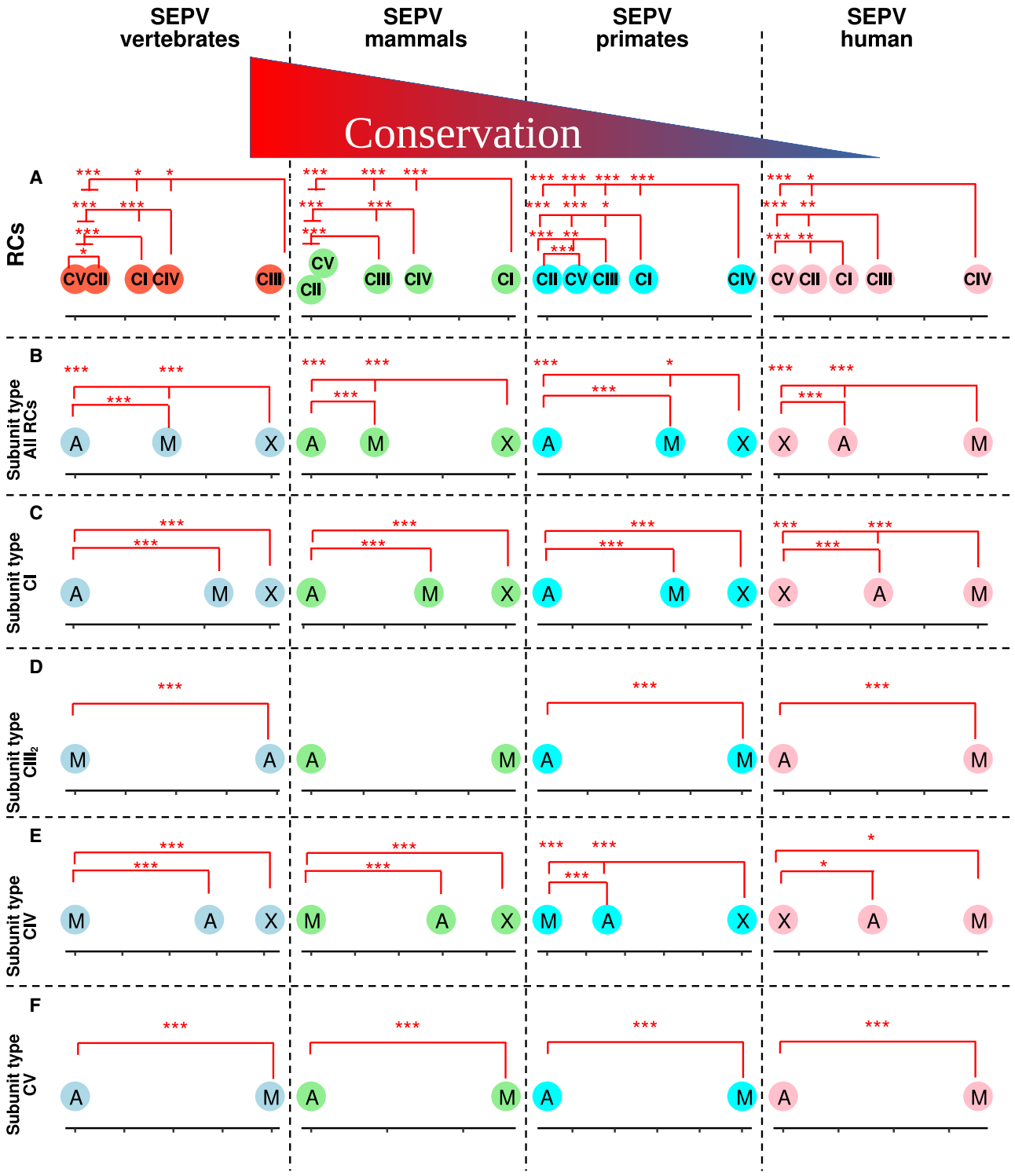


Figure 3. Analysis of evolutionary strategies and population variability in OxPhos RCs

Differences in SEPV are represented in one dimension according to relative values of Kolmogorov-Smirnov distance that separates two consecutive points. In such representation, conservation levels decrease from left to right. Panels depict, from left to right, conservation in vertebrates, mammals, primates, and humans. One-dimensional dot plot comparing conservation by OxPhos RCs (A). Differences in conservation depending on encoding chromosome in all RCs (B),

(legend continued on next page)

points considered except in vertebrates, where it does not differ significantly from the autosomal subunits (Figure 3D). The vertebrate category includes land animals as well as fish and amphibians, for whom oxygen availability differ significantly. This may imply differences in oxygen sensing and the management and production of ROS coupled to the Q-cycle.[14] It is tempting to link this with an increase in the variability associated with CIII.

Conservation by subunit inheritance type in CIV

CIV is the RC that yields electrons to O_2. It is the RC with the largest number of interchangeable isoforms whose presence varies depending on the cell type or physiological or ontological conditions.[15] The functional core of the enzyme is constituted by subunits encoded by the mtDNA (MT-CO1-3). In addition, CIV contains the remaining X-linked subunit in mammals, COX7B. Interestingly, contrary to the general trend observed, from vertebrates to primates, the CIV subunits encoded by the mtDNA are conserved the most, the nuclear-encoded subunits being the main evolutionary driver (Figure 3E). However, in human populations, this trend is reversed, and mtDNA-encoded subunits are the most variable and COX7B (X-linked) the most conserved, likely reflecting the higher mtDNA mutation rate and the Haldane effect, respectively (Figure 3E, right panel). In addition, in humans, no significant differences were observed between autosomal and mtDNA-encoded subunits. These findings, together with the analysis from 1000 Genome Project, reinforce the idea that the CIV presents a unique evolutionary strategy. Here, the variability comes from the nuclear subunits, a fact that is considerably enhanced by the existence of interchangeable isoforms (Figure 2). This strategy will confer to the CIV a greater plasticity for physiological adaptation.

Conservation based on subunit inheritance type in CV

The ATP-synthetase is not part of the ETC, being the enzyme in charge of capitalizing on the H^+ gradient generated by the transport of $NADH/FADH_2$ to O_2. In the CV, from vertebrates to current human populations, the variability of the subunits encoded by the mtDNA is significantly higher (Figure 3F). Considering that CV (and CII) is the most conserved among the evolutionary groups studied, this is probably due to its inherent stoichiometry. Specifically, in its monomeric form (the structurally simplest configuration), CV consists of eight copies of ATP5M1, ATP5MC2, or ATP5MC3 forming the F_1 particle, along with three copies of ATP5F1A and ATP5F1B forming the F_O particle (Figure 1). Since these subunits constitute most of the autosomal component of CV, they are subject to stronger selective pressure, making CV the most conserved complex.

In summary, RCs that participate in supercomplex formation actively pump H^+ across the inner mitochondrial membrane and have mtDNA-encoded subunits as core components that serve as key evolutionary drivers from vertebrates to present-day human populations, particularly CI and CIV. However, CIV exhibited variability largely influenced by nuclear-encoded subunits further amplified by the incorporation of alternative isoforms. Conversely, CV and CII are those under the highest selective pressure. In general, X-linked and especially mtDNA-encoded subunits are the evolutionary drivers in the OxPhos system.

Evolution in action: Conservation of protein-binding sites

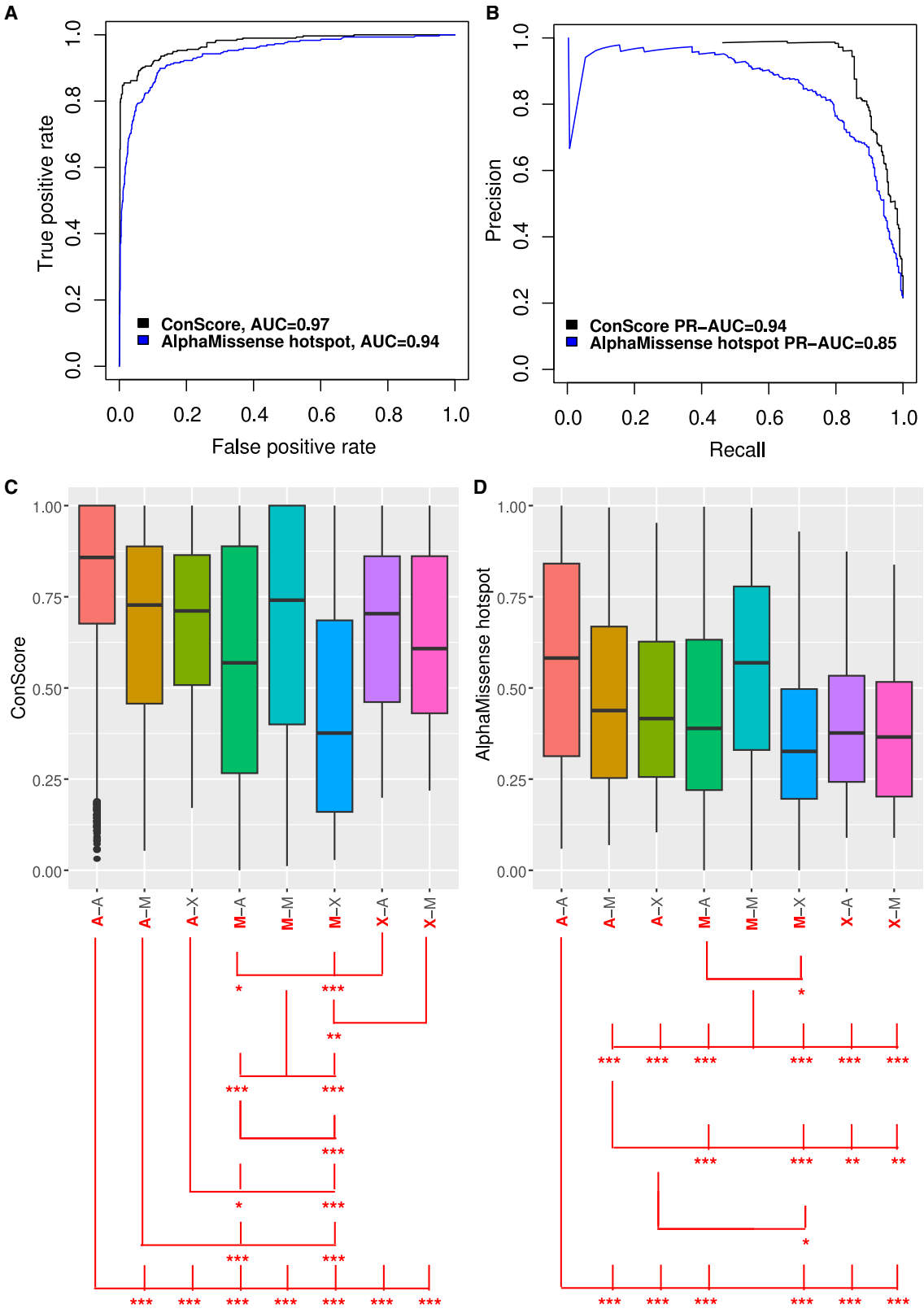
Given that mitonuclear incompatibility is a well-established speciation mechanism and that RCs consist of physically bound proteins, we aim to explore the variability of residues at subunit interface residues (IFRs). We aim to evaluate the evolution of binding sites between interacting subunits in comparison to other regions of the RCs (non-IFRs). This analysis uncovers distinct evolutionary dynamics for IFRs compared to the rest of the protein regions.

The analysis of all subunit types together in vertebrates, mammals, and primates reveals that IFRs exhibit significantly higher conservation than non-IFRs (SEPV levels are lower), with no significant differences among present-day populations (Figure S1A). This pattern of higher conservation of IFRs persists when the analysis is restricted to autosomal subunits, extending even to current human populations (Figure S1B). However, in subunits encoded by mtDNA, the trend is reversed (residues at protein interfaces are less conserved than non-IFRs) (Figure S1C). This suggests that selective pressures on IFRs of mtDNA-encoded subunits are relaxed at binding sites with neighboring subunits, allowing for greater evolutionary flexibility. Finally, when analyzing X-linked subunits, we observed the same trend as in autosomal subunits, suggesting a pattern consistent across all nuclear-encoded subunits, except in primates, where the difference is not significant (Figure S1D). The absence of significant differences in primates may indicate a relaxation of evolutionary constraints acting on the IFRs of X-linked subunits in this group. Note that the proportions of IFRs versus non-IFRs were sufficiently representative, constituting more than half of the positions of the subunits encoded by the mtDNA or X-linked subunits, given the structural arrangement of the RCs (Figure S1E).

Next, focusing on IFRs, we investigate whether the conservation at binding sites vary depending on the inheritance mechanisms of the subunits involved. To simplify this analysis, we summarize the evolutionary process by weighting the conservation values within each group into a single novel score, hereafter ConScore (more details in STAR Methods). ConScore gives values from ConScore = 0, non-conserved positions, to ConScore = 1, fixed positions in the evolution; therefore, it follows a different scale than SEPV. Table S1 contains all SEPV values for each subunit position across all phylogenetic groups, as well as the corresponding ConScore values. In addition, ConScore values were depicted in our *in silico* models (Video S1).

The analysis of IFRs by involved subunit type was performed in comparison with AlphaMissense predictions for these positions. AlphaMissense is a machine-learning model that predicts the

in CI (C), in CIII2 (D), in CV (E), and in CIV (F). In all cases, differences in distributions of Shannon entropy are evaluated, depending on the number of compared categories, using the nonparametric Wilcoxon test or Kruskal-Wallis followed by Dunn's *post hoc* test. False discovery rate was corrected by Benjamini-Hochberg method in multiple testing. Comparisons between distributions were performed considering the stoichiometric representation of each subunit within the compared groups. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.



(legend on next page)

pathogenicity of missense variants in human proteins.[16] It uses deep learning integrating evolutionary and structural information to classify these mutations as benign or pathogenic, aiding in the research of genetic diseases. Thus, it serves as a valuable functional predictor, capable of identifying positions that are functionally or structurally relevant.

However, before going into this analysis, and given that AlphaMissense is a functional predictor oriented to the detection of pathogenic mutations, we compare the capability of ConScore with AlphaMissense hotspots for this purpose, considering the missense variants classified in ClinVar[17] that affect the OxPhos system, measured in terms of area under the receiver operating characteristic curve (ROC-AUC) and area under the precision-recall curve (PR-AUC) (Figures 4A and 4B). Although AlphaMissense was trained on a dataset that included these variants, our ConScore outperformed AlphaMissense hotspot results in the detection of pathogenic mutations, presenting an ROC-AUC = 0.97 and AUC-PR = 0.94. Consequently, ConScore was shown to be an efficient measure for weighting the relevance of the OxPhos system positions. Nevertheless, there is a direct strong correlation between ConScore and AlphaMissense hotspots, Spearman's $\rho = 0.726$ ($p < 0.001$) among all OxPhos subunit positions. Table S2 includes the labeled ClinVar dataset used for benchmarking ConScore.

Considering the types of subunits that define the binding site, A-A homologous interfaces are the most conserved. In contrast, the heterologous IFRs of mtDNA-encoded subunits interacting with nuclear subunits (and especially those binding to X-linked subunits) exhibit the highest variability. This suggests that the relaxation of selective pressures on IFRs of mtDNA-encoded subunits primarily occurs at heterologous junctions (M-A and M-X), enhancing their adaptability to nuclear partners (Figure 4C). Following the same reasoning, selective pressures on IFRs in autosomal subunits are also relaxed when they form heterologous binding sites, interacting with X-linked or mtDNA-encoded subunits. Although it does not directly measure conservation, AlphaMissense hotspot perfectly captures homologous/heterologous divergence of residues involved in autosomal subunit binding (except against binding sites between mtDNA-encoded subunits, M-M) (Figure 4D).

In summary, IFRs are more conserved than non-IFRs, except in mtDNA-encoded subunits, where the opposite pattern is observed. Considering the potential binding partners, junctions between autosomal subunits (the most represented group with 9,417 IFRs) are also the most conserved, likely to prevent mismatches associated with their diploid nature. In contrast, IFRs from mtDNA-encoded subunits involved in heterologous interactions (particularly M-A junctions, comprising 1,922 residues, the second-largest IFR group) experience the second most relaxed selective pressures. This relaxation may facilitate mitonuclear compatibility, given the variability introduced by diploidy and heterozygosity, not only in present-day human populations but also across different evolutionary lineages.

Relationship between population variability, level of conservation, and impact on structure

Next, we aimed to explore whether the two-conservation metrics, ConScore (a weighted measure of conservation since vertebrates) and human population constraints (SEPV in humans), were linked to RCs structural impact. Here, structural impact is defined as the resilience to energy fluctuations at each position when mutated into any of the 19 possible amino acid alternatives, a metric we refer to as probability of structural impact (PSI) (see STAR Methods for details). However, since more than 25% of the positions are fixed in the evolution (ConScore = 1), the stratification of the analysis was performed according to the conservatism score bands defined by the first quartile (Q1; $n = 5,300$), the second quartile (Q2; $n = 5,302$), non-fixed positions belonging to the third quartile (Q3 non-fixed; $n = 3,857$), and the fixed positions in the evolution (fixed; $n = 6,739$) (Figure S2A). As expected, there was clear direct association between the degree of conservation and the PSI. However, when we quantified these differences, measuring the size effect as the Cohen's d, we observed that, although significant, they were considerably small (Q1 vs. Q2, d = 0.09, 95% confidence interval [CI95] = [0.05–0.13]; Q2 vs. Q3 non-fixed, d = 0.19, CI95 = [0.15–0.23]; and Q3 non-fixed vs. fixed, d = 0.12, CI95 = [0.08–0.16]). This makes sense if we consider that the relevance of a residue in a protein is not always structural and, on the contrary, depending on where a mutation occurs, it can have a high structural impact without affecting the function of the complex.

Regarding human populations, approximately 68% of OxPhos positions are fixed, making it impossible to analyze structural impact by quartiles of SEPV. Instead, we categorized positions based on whether they were fixed (SEPV = 0, $n = 14,527$) or variable (SEPV > 0, $n = 6,671$) (Figure S2B). Although we found significant differences in PSI between fixed and non-fixed positions, the effect size was minimal (d = 0.07, CI95 [0.04–0.10]). These results likely reflect two key aspects of our data: first, that, while gnomAD v3 is one of the largest human-variability databases, it does not comprehensively capture all variation; second, as previously discussed, residue relevance does not necessarily correlate with structural impact.

Since evolutionary differences were observed depending on the RC and subunit type, as discussed in previous sections,

Figure 4. Evaluation of ConScore and applied analysis of binding-site conservation based on the combination of subunit types

(A and B) Comparison of ConScore benchmarking with AlphaMissense hotspots, assessed based on performance in the ROC curve (A) and the PR curve (B). (C and D) Boxplot depicting differences between binding sites depending on the subunit types involved in the junction for ConScore values (C) and AlphaMissene values (D). Precision-recall curves for ConScore and AlphaMissense. Note that the ConScore curve starts at a recall of approximately 0.5 because a large proportion of variants receive the same maximum score of 1.0, leading to a "tie" at the top of the ranked list. This results in the initial retrieval of more than 50% of the true positive variants at once. Differences in the distributions of ConScore were assessed based on the number of categories compared.

Statistical analyses were conducted using the Kruskal-Walli's test followed by Dunn's *post hoc* test. Multiple testing corrections were applied using the Benjamini-Hochberg method to control the false discovery rate. Comparisons between distributions considered the stoichiometric representation of each subunit within the analyzed groups. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

we aimed to further investigate whether these factors also influence structural impact. By evaluating potential variations in structural impact across different RCs and subunit types, we sought to determine whether the same evolutionary patterns extend to the structural level or whether distinct trends emerge in relation to these conditions. Considering the different types of subunits, those encoded by mtDNA are susceptible to greater structural impact (Figure S2C). This makes sense given that these subunits occupy central positions in most RCs, constituting the core of the RCs (in CI, CIII and CIV). However, these differences, though significant, are low between mtDNA-encoded and autosomal subunits (d = 0.39, CI95 = [0.35–0.42]) or moderate between mtDNA-encoded and X-linked subunits (d = 0.48 CI95 = [0.36–0.61]). This is because the autosomal subunits may also constitute the core of the RCs (CII, CV, or N and Q modules in CI), whereas the X-linked subunits occupy more peripheral positions in the structures. We also found differences in PSI among all the RCs (Figure S2D). However, the differences in all cases were small (Cohen's d <0.5) (Table S3). On the other hand, presenting a low size effect that, under normal physiological conditions, does not represent a problem, they could become significant in extreme conditions, such as in case of infections where the ability of an individual to withstand higher levels of hyperthermia may be associated with small differences in structural stability. Table S4 contains all PSI values for each subunit position regarding our *in silico* models.

In summary, we observed a direct association between conservation, including human population, and structural resilience, but this association is weak. These findings apply also to differential levels of conservation as RCs or subunit types. Thus, sequence conservation is primarily associated with the specific functional roles of a protein region, such as active sites or ligand-binding domains, whereas structural stability is influenced by broader, global interactions between residues. In consequence, in RCs, regions with lower sequence conservation can still harbor low structural resilience even if they are not directly essential for function.

Extreme constriction of genetic variability in OxPhos structural proteins

Thus far, we have attempted to define the landscape of OxPhos variability through vertebrate evolution to the present day. However, this map would be incomplete without the analysis of extreme strategies for shaping human population constraints in OxPhos system. Here, we aim to determine two extreme strategies that influence variability by directly impacting the heterozygous state of autosomal genes, namely homozygosity bias and its most extreme manifestation, the fixation of a residue (SEPV = 0) within the structure through human populations. The primary goal of this analysis is to identify sites where selective pressures vary in intensity, allowing us to define the most constrained positions based on human population data.

We first explore, in those cases where population variability is allowed, whether the heterozygosity frequency is tolerated or penalized. We found 11 homozygous biased positions (Table S5). These positions indicate a layer of population-level restriction less stringent than absolute residue fixation for autosomal subunits despite its diploid condition. Perhaps, in a larger sample of human variability, it would be possible to detect a larger number of constrained positions by this mechanism. However, these findings are relevant because they show that, at least for these 11 positions, there is a sweeping against variability at individual level.

When we analyzed completely restricted positions, we estimate that, of the 26,623 residue positions in all the human OxPhos complexes, 18,208 have no described amino acid variability, representing 64.5%–74% of residues depending on the RC and thus revealing high conservation of the OxPhos system in humans. First, we examined whether fixed positions were randomly distributed across RCs. Our analysis revealed that CV and CII, the most evolutionarily conserved RCs, were significantly enriched in fixed positions. In contrast, CI, CIII, and CIV exhibited an underrepresentation of them, aligning with previously observed trends (Table S6). CII was the only RC without a biased distribution of fixed positions. Fixed positions were projected in our 3D structures by RC and subunit type (Video S1).

Next, we conducted an enrichment analysis of fixed positions within OxPhos subunits (Figure 5; Table S7). Interestingly, despite mtDNA-encoded subunits being the most variable, some (MT-ND4, MT-ND4L, and MT-CO1) showed a significant enrichment of fixed positions. As expected, the X-linked subunits NDUFA1 and NDUFB11 were also enriched in fixed positions. In contrast, for some of the CIV isoforms, the number of fixed positions was lower than expected, as is the case for COX6A1 and COX6A2 (which are mutually exclusive), COX7A2L (a key isoform involved in the formation of the CIV-$CIII_2$ supercomplex), and COX8C. This further supports the idea that CIV's major source of variability is the assembly of interchangeable isoforms.

Finally, it is noteworthy that subunits repeated within the structure of CV were primarily enriched in fixed positions. This reinforces the notion that the stoichiometric balance of CV plays a key role in its high level of conservation.

Allelic imbalance as a downstream mechanism to limit the potential variability in autosomal OxPhos genes

Allelic imbalance (AIM) is a phenomenon by which the two alleles for the same gene in a heterozygotic context are expressed differentially. This is a regulated feature in response to genetic or epigenetic factors. In consequence, while selection and genetic drift determine potential availability, AIM can reshape the effective genetic landscape. Thus, we explore whether this mechanism influenced the effective OxPhos variability finally assembled. AIM can vary depending on the tissue or even at single-cell level. In this study, we analyzed single-cell RNA-seq data from different sources to evaluate the existence or not of AIM in OxPhos genes at the single-cell level. We used binomial test per gene and per cell to assess AIM (adjusted p <0.05).

We analyzed samples from two mouse models that belong to different subspecies, the inbred strain C57BL6/J from *Mus musculus musculus* and *Mus musculus castaneus*. These subspecies separated about 300,000–500,000 years ago and consequently have genome-wide differences in the order of millions of variants. Since these subspecies represent an evolutionary point prior to speciation still capable of generating fertile progeny, hybrids from alternative crosses between them should shed light

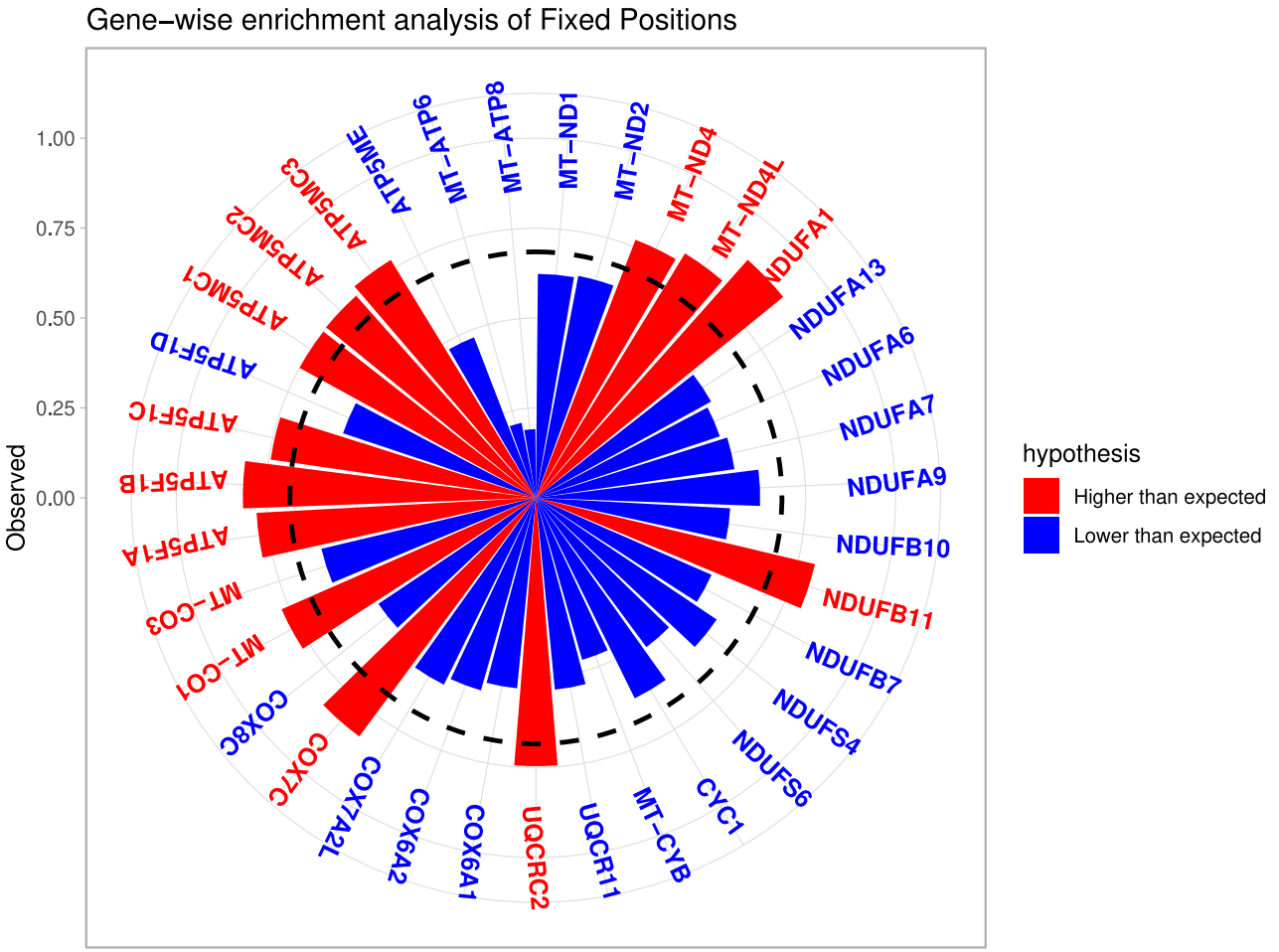


Figure 5. Gene-wise enrichment analysis in fixed positions

In all cases, null hypothesis of gene-wise random distribution of fixed positions was evaluated applying a binomial test. Only genes with biased distributions were represented, considering a threshold for adjusted $p < 0.05$ corrected by the Benjamini-Hochberg method.

on AIM patterns in OxPhos genes. We first analyzed single-cell RNA-seq data in adult fibroblasts from different individuals, from different sexes and alternative crosses (Figure 6). As expected, mtDNA and X-linked OxPhos genes were found to be monoallelic. Strikingly, AIM was observed in autosomal genes, which were monoallelic in favor of expression of the C57 allele, irrespective of the sex or cross of the individuals from which the cells were coming. This was the case of the OxPhos genes *Ndufb2*, *Ndufs5*, *Ndufs6*, and *Ndufa12* in CI; *Sdhd* in CII; *Uqcrh* and *Uqcr10* in $CIII_2$; *Cox5b*, *Cox7a2*, *Cox7a2l*, *Cox7c*, and *Cox6c* in CIV; *Atp5fe1*, and *Atp5pb*, *Atp5mc1*, *Atp5pd*, and *Atp5mk* in CV. To our surprise, cytochrome *c* (*Cycs*), responsible for transferring electrons from $CIII_2$ to CIV, also showed AIM in favor of C57.

We also analyzed embryonic cells from different embryonic tissues at the E15 stage, showing AIM expression in favor of the C57 allele for several autosomal OxPhos genes, although the number of analyzed cells was low (four brain cells, 31 liver cells, and 38 fibroblasts) (Figure S3A). AIM toward the C57 allele was shown in *Ndufa6*, *Ndufc1*, *Ndufs5*, *Ndufs6*, *Ndufa12*, *Sdha*, *Sdhd*, *Uqcr10*, *Cox5b*, *Cox6c*, *Cox7a2*, *Cox7a2l*, *Cox7c*, and *Cox6b1*, as well as *Atp5f1e*, *Atp5pd*, *Atp5mc1*, *Atp5pb*, *Atp5mk*, and *Cycs*. Except for some genes, such as *Ndufc1* or *Cox6b1*, most of these genes maintain the C57 AIM direction observed in adult fibroblasts. This result suggests that AIM is established in an early stage in development and occurs in multiple cell and tissue types.

To determine the developmental stage at which the AIM arise, we assessed AIM in four zygotes, 14 blastomeres obtained from four-cell embryos, 42 blastomeres from eight-cell embryos, and 57 blastomeres from 16-cell morula stage of crosses between a castaneous females and C57 males. As expected, the zygote stage revealed a marked silencing of paternal alleles (Figure S3B). From this initial situation, an early shift in allele expression was observed at the four-cell stage for most of the autosomal OxPhos genes that show C57-skewed AIM in adult fibroblast. This is the case of *Ndufa12*, *Ndufa6*, *Ndufv2* (without significant AIM trend in later stages), *Sdhd*, *Uqcr10*, *Cox5b*, *Cox6c*, *Cox7a2*, *Cox7a2l*, *Atp5pb*, *Atp5pd*, *Atp5mk*, and *Cycs*, while little information was collected for *Ndufs5*. The remaining autosomal

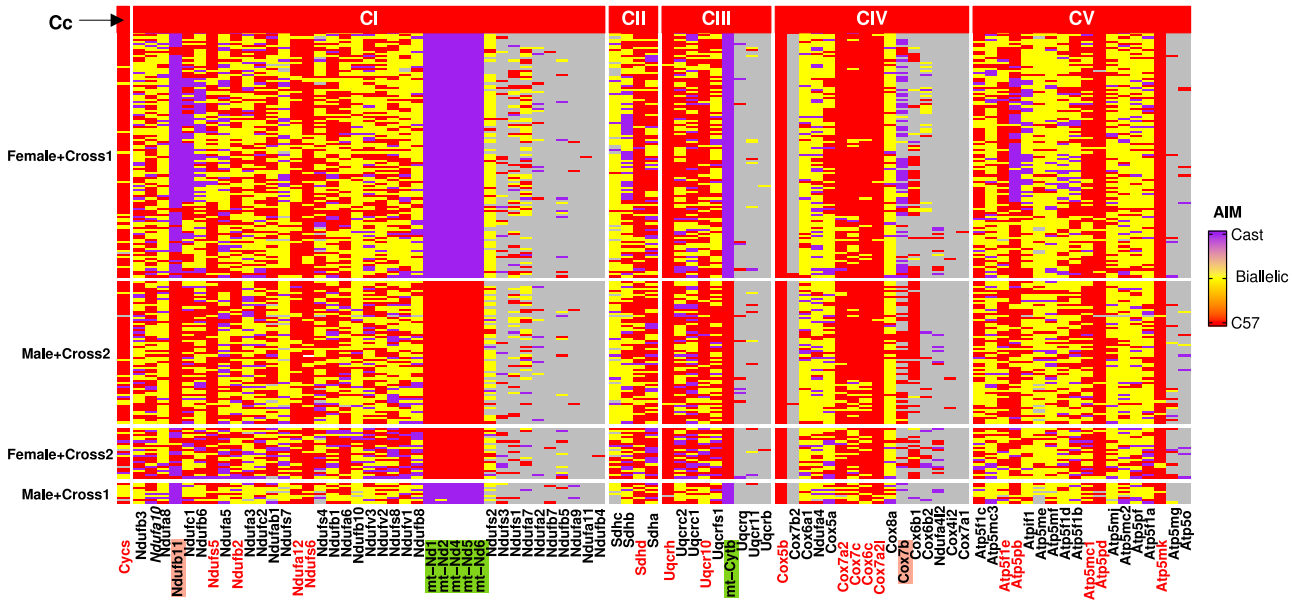


Figure 6. Allelic imbalance in adult fibroblasts from castaneous × C57 complementary crosses

Heatmap representing allele imbalance in OxPhos genes in adult fibroblasts from C57 vs. castaneous (Cast) strain F1 hybrids. Rows represent different cells, clustered by cross type and sex; columns reflect OxPhos genes grouped by RCs. Genes located on mtDNA are highlighted in green squares, those on the X chromosome are highlighted by pink squares. Finally, genes with monoallelic bias toward the C57 allele are shown in red type. AIM toward specific alleles was tested with the binomial test, with the false discovery rate corrected by the Benjamini-Hochberg method, considering a threshold for adjusted $p < 0.05$.

OxPhos genes showed differing degrees of paternal C57 expression, which also varied with the blastomere stage. In contrast, *Ndufa3* showed maternal (castaneous) AIM in these early stages, losing this trend by E15. This analysis shows that, for many genes, C57-biased expression starts very early, in the blastomere stages of development. All results of AIM in OxPhos genes, in all analyzed samples, are gathered in Tables S8–S10, while gene-wise results are in Table S11.

Our data indicate that AIM may serve as an important modifier of the degree of variability encoded in the genome at the individual level. This adds an additional layer of complexity to regulate the fitness of the OxPhos system.

DISCUSSION

The evolutive forces driving the fate of mtDNA since eukaryogenesis are puzzling.[18] In vertebrates, mtDNA still retain 13 protein-coding genes for OxPhos subunits encoded by mtDNA. This unique situation necessitates a co-evolutionary relationship between the mitochondrial and nuclear genomes, incurring significant cellular costs.[19] However, the implications of this mitonuclear association remain incompletely understood.

Previously, other studies have detailed analysis of selective signatures in mtDNA, examining OxPhos genes and identifying positively selected nodal mutations in the phylogeny with implications for evolution and disease within humans.[20] However, our study explores human evolutionary constraints regarding OxPhos proteins as the final stage of this process, offering a 3D context for all mitochondrial subunits (both mtDNA and nuclear encoded). To achieve this, we leverage gnomADv3, a comprehensive human genomic dataset that provides a broad representation of OxPhos diversity across populations, making it a superior resource compared to other datasets.

While mitonuclear interactions have previously been examined at the genomic level in admixed populations,[21] the potential variability at the OxPhos level has yet to be quantified. Our study provides compelling evidence that the diploid condition of autosomal genes, coupled with heterozygosity, represents a crucial source of individual-level variability in the OxPhos system. The coexistence of two mtDNA haplotypes within the same cytoplasm has been shown to trigger robust selection in favor of one mtDNA variant during oogenesis and early embryonic stages.[19,22,23] After birth, some tissues systematically choose one mtDNA haplotype, while other tissues select the other by intracellular selection.[23–25] This type of mtDNA heteroplasmy leads to continuous stress and the risk of organ failure in critical tissues, such as the heart, skeletal muscle, and lungs, which cannot select either mtDNA variant.[3] For some genes, the evolutionary path selected to avoid this problem has been to limit expression to a single allele, a pattern observed in genes encoded by mtDNA and the X chromosome (e.g., Ndufa1, Ndufb11, and Cox7b, which are also encoded by X chromosome in placental mammals). However, our SEPV analysis of human populations shows that mtDNA subunit variability in human populations far exceeds that of nuclear subunits. This suggests that evolution has preserved a double strategy for these 13 mitochondrial subunits, emphasizing high population-level variability with minimal individual-level variability.

Conversely, genes fixed on the X chromosome, which are also monoallelicaly expressed, exhibited lower human population variability, possibly due to selective sweeping in males (hemizygotes for X-linked genes) while adhering to different rules of

sexual transmission and sister chromatid recombination in females. These results align with previous studies showing that X chromosome diversity is lower than that of autosomes.[26–28] In addition, evolutionary findings highlighted that each chromosome/inheritance mode has different effects depending on the RC.

The comparative analysis of conservation from vertebrates to primates at single-residue resolution highlights the singular contribution of the X-encoded subunits. Among human populations, diversity is low, being a major source of interspecies variability, because of the Haldane effect, which in the long-term causes X-linked genes to evolve faster than autosomal genes. This highlights the potential role of X-encoded subunits in speciation by driving nucleus-mitochondria compatibility.

The higher mutational rate of mtDNA represents a constant source of change through evolution. Preservation of these 13 genes in mtDNA, with a high mutation rate and uniparentally transmitted, allows population-level adaptation to the high diversity that autosomal genes represent at the individual level. This would indicate that the higher mutational rate of mtDNA has evolved to allow compatible combinations of OxPhos system components, after passing the various selective checkpoints. The higher mutational rate of mtDNA should then be considered a selected feature rather than the consequence of its location in a mutagenic environment or due to the lack of repairing mechanisms. This aligns with the Deem and Earl idea proposing evolvability as an evolutionarily selectable trait.[29]

The difference between interspecific conservation and that observed among human populations arises from the general decline in variability due to the fixation of species-specific residues. However, mtDNA, with its inherently higher mutation rate, remains the primary source of intraspecific variability, driving genetic diversity within the same species. This estimation supports Wallace's 2010 hypothesis that a high mutation rate continually introduces nonlethal mtDNA variants into the general population, disrupting mitochondrial bioenergetics and facilitating species' adaptation to their environment.[30] Interestingly, in this study, we demonstrate that this role varies depending on the RC.

CI and CIV showed the greatest divergence between species. For CI, mtDNA is the main agent of change, whereas for CIV, change is driven by subunits encoded in the nucleus, enhanced by alternative isoforms. This conclusion has experimental support by the generation of interspecific xenomitochondrial cybrids harboring human nucleus and mtDNA from non-human apes.[31] In these cells, CII, CIII, CIV, and CV showed normal activity, while CI function was deficient in all interspecific crosses.[31] By the same token, xenomitochondrial cybrids between mouse nucleus and rat mtDNA showed concomitant impairment of CI and CIV.[32] An additional difference between CI and CIV variability involves the effect of diploidy + heterozygosity. This is marginal for CIV, which has instead evolved the *ad hoc* strategy of using alternative paralogous isoforms. Thus, in CIV nuclear-encoded subunits, autosomal and X-linked COX7B, not mtDNA, are the main drivers of evolutionary changes. On the other side of the scale, CV and CII are the most conserved RCs, due to their stoichiometric composition (CV) and to the lack of mtDNA-encoded or X-linked subunits (CII).

Alternative to our proposal, the observed variability might result from genetic drift after the speciation process. We found this explanation unlikely, as evaluating the collective behavior of a group of residues can reduce errors in interpreting their role as drivers of speciation. We observe consistent patterns that are non-random and concentrated within residues of CI and CIV or mtDNA. This strongly reduces the likelihood of error in asserting their role in driving speciation, as genetic drift would not produce a defined pattern.

The novel conservation score, ConScore, summarizes and weights the conservation for each subunit position in four nested phylogenetic groups. Although this metric was not specifically designed for clinical uses, its ability to distinguish positions associated with pathogenic and neutral mutations in the OxPhos system proved to be highly effective, outperforming AlphaMissense hotspots, a tool designed to identify potentially deleterious mutations. Therefore, we propose ConScore as a valuable tool for prioritizing missense mutations in the OxPhos system for clinical applications. Nonetheless, both tools exhibited a strong correlation of their values across all OxPhos positions. ConScore is a tool designed to account for human population constraints by incorporating three additional nested groups, providing a balanced measure of each position's constraints through the integration of evolutionary information. Consequently, it is useful to perform functional predictions, making it possible to overcome under-sampling among the current human population by borrowing information from other species.

The residues that constitute the inter-subunit binding surface (the IFR) are of particular interest and they should reflect convergence between chromosomal types, especially at heterologous binding sites. MtDNA-encoded and X-linked subunits, being the main evolutionary sources of variability between species, have a differential behavior regarding the residues involved in the binding sites. While IFRs from mtDNA-encoded subunits are more variable than non-IFRs, IFRs from X-linked subunits remain systematically more conserved than non-IFRs (except in primates). Thus, the main evolutionary drivers show divergent strategies of interaction with neighboring subunits, being more conserved in X-linked IFRs than mtDNA-encoded IFRs.

To understand the analysis of the binding sites according to the type of subunits involved and to reduce the amount of analysis needed, we used ConScore in comparison with AlphaMissense hotspots. Again mtDNA-encoded IFRs were revealed as the least conserved ones, especially when they are involved in heterologous binding sites (M-A and M-X). Thus, binding interfaces involving mtDNA-encoded subunits decrease the variability constraints to adapt to their partners (autosomal or X encoded), reinforcing the evolutionary role of mtDNA subunits. Since diploidy plus heterozygosity could increase the level of mismatches, binding sites between autosomal subunits (A-A) were the most conserved ones because they seem to be under higher selective pressures. In general, IFRs at heterologous binding sites exhibit lower restriction levels compared to homologous ones. For instance, M-M interfaces are more conserved than M-A or M-X, just as A-A interfaces are more conserved than A-X or A-M. This relaxation of evolutionary constraints allows adaptation to binding partners with distinct mutation rates, expression patterns (biallelic vs. monoallelic), and inheritance

mechanisms. X-linked IFRs, always involved in heterologous binding sites, exhibited a conservation level like that of autosomal IFRs from heterologous junctions. The fact that interface residues (IFRs) in mtDNA-encoded subunits are less conserved than non-interface residues challenges the conventional theory of protein evolution, which defines binding sites as more conserved areas due to functional constraints.[33] Furthermore, we provide empirical evidence supporting how evolution has acted at heterologous interfaces.

We found a direct, but weak, association between conservation and structural resilience (PSI). This weak association means that conservation is not only linked to structural strength but rather to functional relevance (i.e., to the residues involved in H^+ pumping). It is important to highlight that, in previous studies, 3D structural information has been used to identify compensatory mutations.[34] However, these assessments have been limited to the MT-CO1-3 and MT-CYB proteins within the OxPhos system, evaluating these proteins in isolation without considering their core-encoded partners. Furthermore, detailed molecular-level evolutionary strategies have not been provided for the entire OxPhos system as a complete entity. Our analysis represents a more comprehensive assessment that considers the interplay between nuclear and mitochondrial components as well as a more holistic evolutionary approach.

We also assess the role of AIM as a mechanism to limit the expression of OxPhos variability at the individual level as well as its role in the speciation process. We provide evidence that AIM has a significant influence on the expression of autosomal variants. Although monoallelic expression has been explored in various contexts, its specific role in OxPhos genes and potential impact on mitonuclear compatibility was unexplored. The main mechanism underlying allele-biased expression involves *cis-trans*-regulatory elements, at least in hybrids of the mouse models considered in this study.[35] The existence of AIM toward a specific allele in certain autosomal OxPhos genes limits the expressed variability, complementing direct residue selection. In the mouse model analyzed, F1 crosses between alternative subspecies represent a pre-speciation evolutionary point. The presence of C57-biased AIM in certain OxPhos genes from early developmental stages across different cell types suggests a role for AIM in shaping the OxPhos landscape and speciation. Subsequent missense mutations in monoallelic autosomal OxPhos genes could be a turning point in evolution, leading to species separation due to OxPhos fitness-related incompatibilities.

In summary, our study has shed light on the significant genetic variability present at the individual level in the OxPhos system, primarily due to heterozygosity in diploid genes within the general population. This diversity prompts a significant biological inquiry concerning the possible existence of transcriptional or post-transcriptional mechanisms that restrict the number of complexes variants expressed with respect to those encoded in the genome. These mechanisms could rule gene expression to mitigate the functional implications arising from the significant theoretical heterogeneity observed within the RCs due to the inherent genetic variability of the OxPhos system among humans. However, this individual diversity is severely limited for the subunits encoded by the X chromosome and the mtDNA. MtDNA-encoded subunits exhibit much higher variability at the population level. Additionally, the variability landscape we have defined for the OxPhos system underscores the special importance of the binding interfaces between subunits, especially those encoded by different chromosome types. This map can contribute to the understanding of the portion of human variability associated with an increased risk of diseases in which the OxPhos system plays a crucial role, such as cancer, neurodegenerative diseases, and mitochondrial diseases, and can provide insights into infertility and miscarriage and enhance our understanding of the mechanism of aging.

Limitations of the study

The interpretation obtained from the gnomAD-v3 database may require confirmation with larger population datasets. Extensive analysis of single-cell RNA-seq data based on full-length reads and extensive coverage of cell types and ages will be needed to finally define the role of AIM in shaping the expression of alternative OxPhos subunit alleles. Moreover, additional post-transcriptional regulation mechanisms not investigated here may contribute to reduce the complexity of the expressed RCs.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, José Antonio Enríquez (jaenriquez@cnic.es).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The original code and data have been deposited at https://github.com/GENOXPHOS/Structural_Evolution_OxPhos as https://doi.org/10.5281/zenodo.15207045 and are publicly available as of the date of publication.
- This paper analyzes existing, publicly available data, accessible at:
  - Gene Expression Omnibus under GSE75659 (https://doi.org/10.1038/ng.3678) and GSE45719 (https://doi.org/10.1126/science.1245316).
  - 1000 Genomes Project database (http://www.internationalgenome.org/data/) and gnomAD v3 database (http://gnomad.broadinstitute.org/).
  - AlphaMissense predictions (https://alphamissense.hegelab.org/hotspot).
  - Protein Data Bank (https://www.rcsb.org/).
  - ClinVar Database from NCBI (https://www.ncbi.nlm.nih.gov/clinvar/).
  - Mouse Genome Project (https://www.sanger.ac.uk/data/mouse-genomes-project/).
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

We thank M.M. Muñoz-Hernandez, R. Martínez de Mena, and E.R. Martínez Jiménez for technical assistance. J.A.E. is supported by Human Frotier Science Program (RGP0016/2018), RTI2018-099357-B-I00, PID2021-1279880B, and TED2021-131611B-I00 funded by MCIN/AEI/10.13039/501100011033 and the European Union "NextGenerationEU"/Plan de Recuperación Transformación y Resiliencia/PRTR, CIBERFES (CB16/10/00282). J.L.C.-A. is supported by Fundación "la Caixa" (LCF/PR/HR23/52430010). P.H.-A. is supported by

JDC IJC2020-042679-I. M.R.-M. is supported by a "Severo Ochoa" FPI Fellowship (PRE2021-097721) awarded by MICIU/AEI/10.13039/501100011033 and by European Social Funds (FSE invierte en tu futuro). F.S.-C. is supported by TED2021-131611B-I00 and EQC2021-007294-P funded by MCIN/AEI/10.13039/501100011033 and the European Union "NextGenerationEU"/Plan de Recuperación Transformación y Resiliencia/PRTR and by CIBERCV (CB22/11/00021). L.S.-G. is supported by PTA2020-019067-I MICIU/AEI/10.13039/501100011033.

The CNIC is supported by the Instituto de Salud Carlos III (ISCIII), the Ministerio de Ciencia, Innovación y Universidades (MICIU), and the Pro CNIC Foundation and is a Severo Ochoa Center of Excellence (grant CEX2020-001041-S funded by MICIU/AEI/10.13039/501100011033).

AUTHOR CONTRIBUTIONS

Conceptualization, J.L.C.-A., F.S.-C., and J.A.E.; methodology, J.L.C.-A., M. R.-M., and F.S.-C.; investigation, J.L.C.-A., L.S.-G., M.R.-M., P.H.-A., and F. B.; visualization, J.L.C.-A. and J.A.E.; funding acquisition, F.S.-C. and J.A.E.; project administration, M.C.J.-G.; supervision, F.S.-C. and J.A.E.; writing – original draft, J.L.C.-A., F.S.-C., and J.A.E.; writing – review & editing, all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work, the authors used Perplexity for language and typo correction. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- METHOD DETAILS
  - Analysis of the potential variability of RCs in the 1000 genome project
  - Analysis of population variability and conservation level at a protein residue level as SEPV
  - Determination of fixed positions and homozygous biased positions and enrichment analysis of fixed positions
  - Development of *in silico* models for RCs
  - Conservation score ConScore calculation
  - Interface residues detection
  - ConScore benchmarking compared with AlphaMissense hotspot
  - Calculation of Probability of Structural Impact
  - Assessment of the association between PSI and ConScore, human population constraints, RCs or subunit type
  - Evaluation of allelic imbalance (AIM) in OXPHOS genes at transcriptional level
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xgen.2025.100945.

Received: November 11, 2024
Revised: March 14, 2025
Accepted: June 9, 2025
Published: July 3, 2025

REFERENCES

1. Enriquez, J.A. (2016). Supramolecular Organization of Respiratory Complexes. Annu. Rev. Physiol. *78*, 533–561. https://doi.org/10.1146/annurev-physiol-021115-105031.
2. Latorre-Pellicer, A., Moreno-Loshuertos, R., Lechuga-Vieco, A.V., Sánchez-Cabo, F., Torroja, C., Acín-Pérez, R., Calvo, E., Aix, E., González-Guerra, A., Logan, A., et al. (2016). Mitochondrial and nuclear DNA matching shapes metabolism and healthy ageing. Nature *535*, 561–565. https://doi.org/10.1038/nature18618.
3. Lechuga-Vieco, A.V., Latorre-Pellicer, A., Calvo, E., Torroja, C., Pellico, J., Acín-Pérez, R., García-Gil, M.L., Santos, A., Bagwan, N., Bonzon-Kulichenko, E., et al. (2022). Heteroplasmy of Wild-Type Mitochondrial DNA Variants in Mice Causes Metabolic Heart Disease With Pulmonary Hypertension and Frailty. Circulation *145*, 1084–1101. https://doi.org/10.1161/CIRCULATIONAHA.121.056286.
4. Burton, R.S., Ellison, C.K., and Harrison, J.S. (2006). The sorry state of F2 hybrids: consequences of rapid mitochondrial DNA evolution in allopatric populations. Am. Nat. *168*, S14–S24. https://doi.org/10.1086/509046.
5. Lee, H.Y., Chou, J.Y., Cheong, L., Chang, N.H., Yang, S.Y., and Leu, J.Y. (2008). Incompatibility of nuclear and mitochondrial genomes causes hybrid sterility between two yeast species. Cell *135*, 1065–1073. https://doi.org/10.1016/j.cell.2008.10.047.
6. 1000 Genomes Project Consortium; Auton, A., Brooks, L.D., Durbin, R. M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., and Abecasis, G.R. (2015). A global reference for human genetic variation. Nature *526*, 68–74. https://doi.org/10.1038/nature15393.
7. Chen, S., Francioli, L.C., Goodrich, J.K., Collins, R.L., Kanai, M., Wang, Q., Alföldi, J., Watts, N.A., Vittal, C., Gauthier, L.D., et al. (2024). A genomic mutational constraint map using variation in 76,156 human genomes. Nature *625*, 92–100. https://doi.org/10.1038/s41586-023-06045-0.
8. Conway, P., Tyka, M.D., DiMaio, F., Konerding, D.E., and Baker, D. (2014). Relaxation of backbone bond geometry improves protein energy landscape modeling. Protein Sci. *23*, 47–55. https://doi.org/10.1002/pro.2389.
9. Davis, I.W., and Baker, D. (2009). RosettaLigand docking with full ligand and receptor flexibility. J. Mol. Biol. *385*, 381–392. https://doi.org/10.1016/j.jmb.2008.11.010.
10. Song, Y., DiMaio, F., Wang, R.Y.R., Kim, D., Miles, C., Brunette, T., Thompson, J., and Baker, D. (2013). High-resolution comparative modeling with RosettaCM. Structure *21*, 1735–1742. https://doi.org/10.1016/j.str.2013.08.005.
11. Hernansanz-Agustin, P., and Enriquez, J.A. (2021). Generation of Reactive Oxygen Species by Mitochondria. Antioxidants *10*, 415. https://doi.org/10.3390/antiox10030415.
12. Hernansanz-Agustin, P., Morales-Vidal, C., Calvo, E., Natale, P., Marti-Mateos, Y., Jaroszewicz, S.N., Cabrera-Alarcon, J.L., Acin-Perez, R., Lopez-Montero, I., Vazquez, J., and Enriquez, J.A. (2024). A transmitochondrial sodium gradient controls membrane potential in mammalian mitochondria. Cell *187*, 6599–6613.e21. https://doi.org/10.1016/j.cell.2024.08.045.
13. Crofts, A.R., Hong, S., Ugulava, N., Barquera, B., Gennis, R., Guergova-Kuras, M., and Berry, E.A. (1999). Pathways for proton release during ubihydroquinone oxidation by the bc(1) complex. Proc. Natl. Acad. Sci. USA *96*, 10021–10026. https://doi.org/10.1073/pnas.96.18.10021.
14. Sommer, N., Pak, O., Schörner, S., Derfuss, T., Krug, A., Gnaiger, E., Ghofrani, H.A., Schermuly, R.T., Huckstorf, C., Seeger, W., et al. (2010). Mitochondrial cytochrome redox states and respiration in acute pulmonary oxygen sensing. Eur. Respir. J. *36*, 1056–1066. https://doi.org/10.1183/09031936.00013809.
15. Sinkler, C.A., Kalpage, H., Shay, J., Lee, I., Malek, M.H., Grossman, L.I., and Hüttemann, M. (2017). Tissue- and Condition-Specific Isoforms of Mammalian Cytochrome c Oxidase Subunits: From Function to Human

Disease. Oxid. Med. Cell. Longev. *2017*, 1534056. https://doi.org/10.1155/2017/1534056.

16. Cheng, J., Novati, G., Pan, J., Bycroft, C., Žemgulytė, A., Applebaum, T., Pritzel, A., Wong, L.H., Zielinski, M., Sargeant, T., et al. (2023). Accurate proteome-wide missense variant effect prediction with AlphaMissense. Science *381*, eadg7492. https://doi.org/10.1126/science.adg7492.
17. Landrum, M.J., Lee, J.M., Riley, G.R., Jang, W., Rubinstein, W.S., Church, D.M., and Maglott, D.R. (2014). ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Res. *42*, D980–D985. https://doi.org/10.1093/nar/gkt1113.
18. Hoitzing, H., Gammage, P.A., Haute, L.V., Minczuk, M., Johnston, I.G., and Jones, N.S. (2019). Energetic costs of cellular and therapeutic control of stochastic mitochondrial DNA populations. PLoS Comput. Biol. *15*, e1007023. https://doi.org/10.1371/journal.pcbi.1007023.
19. Latorre-Pellicer, A., Lechuga-Vieco, A.V., Johnston, I.G., Hämäläinen, R. H., Pellico, J., Justo-Méndez, R., Fernández-Toro, J.M., Clavería, C., Guaras, A., Sierra, R., et al. (2019). Regulation of Mother-to-Offspring Transmission of mtDNA Heteroplasmy. Cell Metab. *30*, 1120–1130.e5. https://doi.org/10.1016/j.cmet.2019.09.007.
20. Levin, L., Zhidkov, I., Gurman, Y., Hawlena, H., and Mishmar, D. (2013). Functional recurrent mutations in the human mitochondrial phylogeny: dual roles in evolution and disease. Genome Biol. Evol. *5*, 876–890. https://doi.org/10.1093/gbe/evt058.
21. Zaidi, A.A., and Makova, K.D. (2019). Investigating mitonuclear interactions in human admixed populations. Nat. Ecol. Evol. *3*, 213–222. https://doi.org/10.1038/s41559-018-0766-1.
22. Jenuth, J.P., Peterson, A.C., and Shoubridge, E.A. (1997). Tissue-specific selection for different mtDNA genotypes in heteroplasmic mice. Nat. Genet. *16*, 93–95. https://doi.org/10.1038/ng0597-93.
23. Sharpley, M.S., Marciniak, C., Eckel-Mahan, K., McManus, M., Crimi, M., Waymire, K., Lin, C.S., Masubuchi, S., Friend, N., Koike, M., et al. (2012). Heteroplasmy of mouse mtDNA is genetically unstable and results in altered behavior and cognition. Cell *151*, 333–343. https://doi.org/10.1016/j.cell.2012.09.004.
24. Lechuga-Vieco, A.V., Latorre-Pellicer, A., Johnston, I.G., Prota, G., Gileadi, U., Justo-Méndez, R., Acín-Pérez, R., Martínez-de-Mena, R., Fernández-Toro, J.M., Jimenez-Blasco, D., et al. (2020). Cell identity and nucleo-mitochondrial genetic context modulate OXPHOS performance and determine somatic heteroplasmy dynamics. Sci. Adv. *6*, eaba5345. https://doi.org/10.1126/sciadv.aba5345.
25. Tostes, K., Dos Santos, A.C., Alves, L.O., Bechara, L.R.G., Marascalchi, R., Macabelli, C.H., Grejo, M.P., Festuccia, W.T., Gottlieb, R.A., Ferreira, J.C.B., and Chiaratti, M.R. (2022). Autophagy deficiency abolishes liver mitochondrial DNA segregation. Autophagy *18*, 2397–2408. https://doi.org/10.1080/15548627.2022.2038501.
26. Arbiza, L., Gottipati, S., Siepel, A., and Keinan, A. (2014). Contrasting X-linked and autosomal diversity across 14 human populations. Am. J. Hum. Genet. *94*, 827–844. https://doi.org/10.1016/j.ajhg.2014.04.011.
27. Hammer, M.F., Woerner, A.E., Mendez, F.L., Watkins, J.C., Cox, M.P., and Wall, J.D. (2010). The ratio of human X chromosome to autosome diversity is positively correlated with genetic distance from genes. Nat. Genet. *42*, 830–831. https://doi.org/10.1038/ng.651.
28. Xu, K., Oh, S., Park, T., Presgraves, D.C., and Yi, S.V. (2012). Lineage-specific variation in slow- and fast-X evolution in primates. Evolution *66*, 1751–1761. https://doi.org/10.1111/j.1558-5646.2011.01556.x.
29. Earl, D.J., and Deem, M.W. (2004). Evolvability is a selectable trait. Proc. Natl. Acad. Sci. USA *101*, 11531–11536. https://doi.org/10.1073/pnas.0404656101.
30. Wallace, D.C. (2010). Mitochondrial DNA mutations in disease and aging. Environ. Mol. Mutagen. *51*, 440–450. https://doi.org/10.1002/em.20586.
31. Barrientos, A., Kenyon, L., and Moraes, C.T. (1998). Human xenomitochondrial cybrids. Cellular models of mitochondrial complex I deficiency. J. Biol. Chem. *273*, 14210–14217. https://doi.org/10.1074/jbc.273.23.14210.
32. Dey, R., Barrientos, A., and Moraes, C.T. (2000). Functional constraints of nuclear-mitochondrial DNA interactions in xenomitochondrial rodent cell lines. J. Biol. Chem. *275*, 31520–31527. https://doi.org/10.1074/jbc.M004053200.
33. Mintseris, J., and Weng, Z. (2005). Structure, function, and evolution of transient and obligate protein-protein interactions. Proc. Natl. Acad. Sci. USA *102*, 10930–10935. https://doi.org/10.1073/pnas.0502667102.
34. Levin, L., and Mishmar, D. (2017). The genomic landscape of evolutionary convergence in mammals, birds and reptiles. Nat. Ecol. Evol. *1*, 41. https://doi.org/10.1038/s41559-016-0041.
35. Goncalves, A., Leigh-Brown, S., Thybert, D., Stefflova, K., Turro, E., Flicek, P., Brazma, A., Odom, D.T., and Marioni, J.C. (2012). Extensive compensatory cis-trans regulation in the evolution of mouse gene expression. Genome Res. *22*, 2376–2384. https://doi.org/10.1101/gr.142281.112.
36. Reinius, B., Mold, J.E., Ramsköld, D., Deng, Q., Johnsson, P., Michaëlsson, J., Frisén, J., and Sandberg, R. (2016). Analysis of allelic expression patterns in clonal somatic cells by single-cell RNA-seq. Nat. Genet. *48*, 1430–1435. https://doi.org/10.1038/ng.3678.
37. Deng, Q., Ramsköld, D., Reinius, B., and Sandberg, R. (2014). Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. Science *343*, 193–196. https://doi.org/10.1126/science.1245316.
38. Keane, T.M., Goodstadt, L., Danecek, P., White, M.A., Wong, K., Yalcin, B., Heger, A., Agam, A., Slater, G., Goodson, M., et al. (2011). Mouse genomic variation and its effect on phenotypes and gene regulation. Nature *477*, 289–294. https://doi.org/10.1038/nature10413.
39. McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P., and Cunningham, F. (2016). The Ensembl Variant Effect Predictor. Genome Biol. *17*, 122. https://doi.org/10.1186/s13059-016-0974-4.
40. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.
41. Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797. https://doi.org/10.1093/nar/gkh340.
42. Grant, B.J., Rodrigues, A.P.C., ElSawy, K.M., McCammon, J.A., and Caves, L.S.D. (2006). Bio3d: an R package for the comparative analysis of protein structures. Bioinformatics *22*, 2695–2696. https://doi.org/10.1093/bioinformatics/btl461.
43. Chaudhury, S., Lyskov, S., and Gray, J.J. (2010). PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta. Bioinformatics *26*, 689–691. https://doi.org/10.1093/bioinformatics/btq007.
44. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.
45. UniProt Consortium (2025). UniProt: the Universal Protein Knowledgebase in 2025. Nucleic Acids Res. *53*, D609–D617. https://doi.org/10.1093/nar/gkae1010.
46. Guo, R., Zong, S., Wu, M., Gu, J., and Yang, M. (2017). Architecture of Human Mitochondrial Respiratory Megacomplex I(2)III(2)IV(2). Cell *170*, 1247–1257.e12. https://doi.org/10.1016/j.cell.2017.07.050.
47. Zong, S., Wu, M., Gu, J., Liu, T., Guo, R., and Yang, M. (2018). Structure of the intact 14-subunit human cytochrome c oxidase. Cell Res. *28*, 1026–1034. https://doi.org/10.1038/s41422-018-0071-1.

48. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. GigaScience *10*, giab008. https://doi.org/10.1093/gigascience/giab008.

49. Saito, T., and Rehmsmeier, M. (2017). Precrec: fast and accurate precision-recall and ROC curve calculations in R. Bioinformatics *33*, 145–147. https://doi.org/10.1093/bioinformatics/btw570.

50. Sing, T., Sander, O., Beerenwinkel, N., and Lengauer, T. (2005). ROCR: visualizing classifier performance in R. Bioinformatics *21*, 3940–3941. https://doi.org/10.1093/bioinformatics/bti623.

51. Moon, C.P., and Fleming, K.G. (2011). Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers. Proc. Natl. Acad. Sci. USA *108*, 10174–10177. https://doi.org/10.1073/pnas.1103979108.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Allele-sensitive single-cell RNA-seq from primary mouse fibroblasts	GEO (Gene Expression Omnibus)[36]	GSE75659
Allele-sensitive single-cell RNA-seq from first generation mouse strain crosses from zygote to the cells of the late blastocyst	Deng et al., 2014; GEO (Gene Expression Omnibus)[37]	GSE45719
5XTD	RCSB PDB	https://doi.org/10.2210/pdb5XTD/pdb
5XTE	RCSB PDB	https://doi.org/10.2210/pdb5XTE/pdb
5Z62	RCSB PDB	https://doi.org/10.2210/pdb5Z62/pdb
1ZOY	RCSB PDB	https://doi.org/10.2210/pdb1ZOY/pdb
4YTP	RCSB PDB	https://doi.org/10.2210/pdb4YTP/pdb
6ZBB	RCSB PDB	https://doi.org/10.2210/pdb6ZBB/pdb
6J54	RCSB PDB	https://doi.org/10.2210/pdb6J54/pdb
6TT7	RCSB PDB	https://doi.org/10.2210/pdb6TT7/pdb
6ZMR	RCSB PDB	https://doi.org/10.2210/pdb6ZMR/pdb
6Z1R	RCSB PDB	https://doi.org/10.2210/pdb6Z1R/pdb
2CLY	RCSB PDB	https://doi.org/10.2210/pdb2CLY/pdb
2OS8	RCSB PDB	https://doi.org/10.2210/pdb2OS8/pdb
6J5I	RCSB PDB	https://doi.org/10.2210/pdb6J5I/pdb
1VZS	RCSB PDB	https://doi.org/10.2210/pdb1VZS/pdb
1BMF	RCSB PDB	https://doi.org/10.2210/pdb1BMF/pdb
2W6E	RCSB PDB	https://doi.org/10.2210/pdb2W6E/pdb
2XND	RCSB PDB	https://doi.org/10.2210/pdb2XND/pdb
2W6H	RCSB PDB	https://doi.org/10.2210/pdb2W6H/pdb
1MAB	RCSB PDB	https://doi.org/10.2210/pdb1MAB/pdb
Other		
1,000 Genomes Project database	The 1000 Genomes Project Consortium[6]	http://www.internationalgenome.org/data/
gnomAD v3 database	gnomAD Consortium[7]	http://gnomad.broadinstitute.org/
AlphaMissense predictions (hotspots)	AlphaMissense hotspots[16]	https://alphamissense.hegelab.org/hotspot
ClinVar database	NCBI[17]	https://www.ncbi.nlm.nih.gov/clinvar/
SNPs from the Mouse Genome Project	Sanger institute[38]	https://www.sanger.ac.uk/data/mouse-genomes-project/
Software and Algorithms		
Variant Effect Predictor (VEP)	Ensembl, EMBL-EBI[39]	https://www.ensembl.org/info/docs/tools/vep/index.html
BLASTP	NCBI[40]	https://blast.ncbi.nlm.nih.gov/Blast.cgi
MUSCLE	Robert C. Edgar[41]	https://www.drive5.com/muscle/
Rosetta suit of tools	Baker Lab, University of Washington	https://www.rosettacommons.org/
bio3d v-2.4-4	R-package[42]	https://doi.org/10.1093/bioinformatics/btl461
PyRosetta v4.0	Gray Lab, University of Washington[43]	https://www.pyrosetta.org/
PyMol	Schrödinger	http://www.pymol.org/pymol
Picard tools	Broad Institute	http://broadinstitute.github.io/picard
STAR	Dobin et al., 2013[44]	https://github.com/alexdobin/STAR

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
R version 4.4.2	R core team	https://www.r-project.org/
Rstudio	RStudio	https://rstudio.com
All code and analyses that contributed to this work	This paper	https://github.com/GENOXPHOS/Structural_Evolution_OxPhos

METHOD DETAILS

Analysis of the potential variability of RCs in the 1000 genome project

Publicly available data from the 1,000-genome project[6] were analyzed to measure the potential variability in assembled RCs in a human population. Thus, third phase SNPs, that include genotype information from 2,504 individuals were annotated with variant effect predictor VEP[39] and missense variants affecting the OxPhos genes were selected (Figure S4). For this analysis we considered only those missense variants that are going to be part of the final product, therefore we filtered out those affecting to residues that will be removed, such as the initial methionine or the signal peptide for the mitochondrial import in some cases, following Uniprot database annotations.[45] The number of potential combinations of different subunits with respect to the amino acid point changes present per individual was then calculated for each RC of the OxPhos system, considering the existence of repeated subunits and alternative isoforms encoded by paralog genes. For better understanding, described OxPhos alternative paralog genes products for RCs IV and V are schematized in Figures 1G and 1H. Then, calculation of all potential combinations for each RC depends on the number of different subunits with alternative alleles (represented here by differences in amino acid composition), the number of times a subunit is repeated in the structure and the existence of subunits encoded by alternative paralog genes.

1) For CI and CII:

$$C = 2^{xn}$$

2) In the case of $CIII_2$, given that complex III is only available as a dimer:

$$C = 2^{2n}$$

3) In CIV for which different subunits have different alternative isoforms:

$$Subunit_{IV} = 2^{n}_{COX4I1} + 2^{n}_{COX4I2}$$

$$Subun_{VIa} = 2^{n}_{COX6A1} + 2^{n}_{COX6A2}$$

$$Subunit_{VIb} = 2^{n}_{COX6B1} + 2^{n}_{COX6B2}$$

$$Subunit_{VIIa} = 2^{n}_{COX7A1} + 2^{n}_{COX7A2} + 2^{n}_{COX7A2L}$$

$$Subunit_{VIIb} = 2^{n}_{COX7B} + 2^{n}_{COX7B2}$$

$$Subunit_{VIII} = 2^{n}_{COX8A} + 2^{n}_{COX8C}$$

$$Subunit_{NDUFA4} = 2^{n}_{NDUFA4} + 2^{n}_{NDUFA4L2}$$

$$Subunit_{other} = 2^{xn}$$

$$C = Subunit_{other} \times Subunit_{IV} \times Subunit_{VIa} \times Subunits_{VIb} \times Subunit_{VIIa} \times Subunit_{VIIb} \times Subunit_{VIII} \times Subunit_{NDUFA4}$$

4) For CV there ATP5MC1, ATP5MC2 and ATPMC3 are interchangeable isoforms:

$$C = 2^{xn} \times \left(2^{8n}_{ATP5MC1} + 2^{8n}_{ATP5MC2} + 2^{8n}_{ATP5MC3}\right)$$

being n the number of different subunits that compose the RC that has two different alleles and x the number of times a subunit is repeated in the structure.

With this analysis we aim to approximate the potential variability of an example human population, rather than perform an exhaustive computation. Consequently, only those missense variants affecting the canonical transcript were considered. In addition, for the ATP-synthase, the computation of possible combinations is limited to the monomer.

Analysis of population variability and conservation level at a protein residue level as SEPV

Shannon entropy is a concept from information theory is a measurement of the amount of "uncertainty" linked to the possible outcomes of a variable. The Shannon entropy is then calculated using the probabilities of these outcomes as:

$$SEPV = -\sum_{i=1}^{n} F_i \times log(F_i)$$

The Shannon entropy takes positive values, starting from 0, being the larger the higher "uncertainty".

This concept can be transferred to the evaluation of the frequencies of the amino acids described at a particular protein position. From here, if these frequencies come from the distinct residues observed at a particular position in a multiple alignment of sequences (MSA) from different species, we will obtain the degree of disorder of that position, i.e. how variable that protein position is among different species, which is a measure of conservation (Figure S5). In this study, we quantified conservation at the protein position level using multiple sequence alignments (MSA) across three nested phylogenetic groups: vertebrates, mammals, and primates. For vertebrates, we aligned protein sequences from up to 50 species per OxPhos subunit, ensuring balanced representation with approximately one-fifth each from amphibians, birds, fish, reptiles, and mammals. Similarly, for mammals, we aligned sequences from 50 different species, while for primates, MSAs were constructed using protein sequences from 25 different species. Given that the use of sequences from different species, while always within the relevant phylogenetic group, may vary across subunits, thereby artificially introducing variability, we incorporated a correction factor into the SEPV calculation. First each species sequence is weighted regarding the differences compared with the reference sequence, that is the human sequence for the subunit:

$$w_{i\ vs\ human} = \frac{Number\ of\ identical\ items\ against\ human\ sequence}{length\ of\ the\ sequence(including\ gaps)}$$

Next, amino acid frequencies in the column

$$F_{i(corrected)} = \frac{\sum w_{i\ Vs\ human} \cdot I(x_i = aa)}{\sum_{i=1}^{n} w_{i\ Vs\ human}}$$

Where *aa* is each of the different amino acid found in the column, being $I(x_i)=0$ when $x_i \neq aa$ and $I(x_i)=1$ when $x_i=aa$. Then corrected Shannon Entropy Position Variability was calculated:

$$SEPV = -\sum_{i=1}^{n} F_{i(corrected)} \times log\left(F_{i(corrected)}\right)$$

On the other hand, it is possible that the frequencies with which different amino acids occur at a particular protein position refer to the frequency within a population of individuals, in which case the Shannon entropy will be a measure of the degree of variability in that population, at that position, for that protein. In gnomAD-v3, is gathered variant information derived from the analysis of 76,156 genomes including mtDNA information (that was not included in gnomAD-v2 and has a great population bias in gnomAD-v4), allowing the evaluation of OxPhos system. From this database we selected missense variants, that affect OxPhos subunits according to the consequences predicted by variant effect predictor VEP,[39] considering only canonical transcripts. Then, if we use the frequencies derived from the variability collected from the thousands of individuals included in the gnomAD-v.3 database, from different populations, we will obtain an approximation to the variability existing in the human population, that here is called SEPV human.

To handle the same scale for SEPV in these four evolutionary points, min-max normalization was carried out for SEPV vertebrates, SEPV mammals, SEPV primates and SEPV human. Then, SEPV in each evolutionary point takes values from 0 to 1, meaning when SEPV=0, that no alternative amino-acid was found, thus this position is fixed in the group, while when SEPV=1 represents the maximum level of variability among all OxPhos subunits.

Regarding both meanings of Shannon entropy, this measurement is going to summarize two sources of information, the number of different amino acids described in such position and their relative frequencies. In the case of SEPVhumans will represent the degree of variability, or alternatively, the level of constraints within the human population, serving as an intraspecific reflection of ongoing evolution, while SEPV in vertebrates, mammals, or primates, it serves as a direct indicator of conservation, reflecting the outcomes of selective pressures or genetic drift processes.

In the case of SEPV.gnomAD.v3, the higher the value, higher the lever of variability in human population, and in the case of SEPV. inter-species, the larger the value the lower the position is conserved. To handle similar scales for both measurements, in both cases a min-max normalization was performed, considering the distributions along the whole OxPhos system.

To define SEPV in vertebrates, mammals or primates, we gathered reference sequences for OxPhos subunits for ten different species whose genomes are reasonably characterized by using blastp tool[40] against refeseq database. Selected sequence ID for subunit

for considered species are listed in Table S12. In a second step, multiple sequence alignment (MSA) was carried out applying muscle algorithm.[41]

Determination of fixed positions and homozygous biased positions and enrichment analysis of fixed positions

OxPhos protein positions with SEPV human=0 was considered as fixed Positions, not being represented in gnomAD-v3.

To define Homozygous biased positions in autosomal encoded subunits, the observed number of homozygotes reported in gnomAD was compared against the expected number of homozygotes derived from amino acid change frequencies in the population ($\sum p^2_i$), being p_i the frequency of homozygous positions in the population. Therefore, this phenomenon can be tested for the discovery of homozygous biased positions performing a binomial test to assess whether the number of observed homozygotes at a certain protein position is larger than expected by mere allele frequency, considering as significant threshold of $p<0.05$ and a false discovery rate adjusted corrected by Benjamini-Hochberg.

To explore whether the number of fixed positions was higher or lower than expected by random distribution by RC or by subunit, binomial test was performed. False discovery rate was corrected applying Benjamini-Hochberg method. We adopted a threshold of significance of adjusted p-value$\leq$0.05.

Development of *in silico* models for RCs

In silico modelling of RCs could be used to integrate population information with the physical interaction between subunits. This specific application is specially to study evolutionary interaction between subunits with different inheritance rules encoded by autosomal, X or mitochondrial chromosome. Although the structures of most human RCs have already been published, these structures need substantial further refinement to be used. For Complex V, no structure is yet available. For structure refinement or for whole structure development, it was used ROSETTA biomolecular modelling suite of tools. Rosetta is a software for protein structure prediction and functional design, that explore randomly conformational space for a protein and test the model in energy terms, using knowledge-based energy functions. Therefore, during the modeling, Rosetta produces several models guided by energy functions, as part of the strategy of increasing the probability of obtaining a model close to the native protein. The strategy for the development of *in silico* models is necessarily different depending on the existence of available structures. In silico models for CI, CIII and CIV, were developed by refinement of structures 5xtd, 5xte and 5z62 respectively.[46,47] These models were obtained by cryo-electron-microscopy and regions with high mobility (i.e., loops) usually remains unsolved with this technique. On top of this, all these models presented several clashes (atoms overlapping in the space), with lot of them involves atoms from the backbone, consequently, given the energetic implications, they cannot occupy those positions in space. Therefore, these models must be refined. To that end, we produced 100 models applying the Rosetta relax protocol.[8] then the most representative model was selected as the lowest energy model within the most crowded cluster. To this aim, unsolved regions of each subunit were modeled independently. In this task, 100 models were produced by Rosetta comparative modeling protocol,[10] using the relaxed incomplete subunit as template. Once again, the most representative model (lowest energy model from the most crowded cluster) of each of these subunits was selected and placed together with the other subunits in the multimeric structure and the whole complex went through an energy minimization (relax protocol) step, producing 100 models.

For complexes CII and CV there are not resolved PDB models for human, but there are resolved structures from other species. Then given the high degree of homology between species for OxPhos proteins, each subunit was obtained by homology modeling (RosettaCM). Here we generated in each case 1,000 models of each subunit using PDB structures from other species as templates (Tables S13 and S14). Then, taking advantage on the high degree of homology for OxPhos subunits between species, all subunits were aligned against most homologous template used in the homology modeling step. Composed structures were relaxed, and the selected models were docked for their prosthetic groups using Rosetta ligand docking tool[9] obtaining 1,000 models for each ligand. In this step most suitable models were selected directly as the lowest energy ones at ligand-pocket interface.

In all cases the visualization of the pdb structure and the video were performed using PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

Conservation score ConScore calculation

For the development of the conservation score, **ConScore** for each position of each subunit of the OxPhos system, the conservation level in the four groups analyzed (vertebrates, mammals, primates and humans) were weighted according to the distribution of fixed positions and overall conservation level within the OxPhos system (Figure S6). Consequently, the conservation values measured by SEPV in each evolutionary point were rank normalized and summarized according to the following expression:

$$x = \sum_{ep\,=\,vertebrates}^{ep\,=\,human} \frac{rank\left(SEPV_{ep}\right)}{N}$$

where $SEPV_{ep}$ represents the protein position conservation values at each evolutionary point, and N denotes the total number of OxPhos positions, considering all distinct compounding subunits.

Subsequently, the rank-normalized conservation values were summarized and transformed to range from 0 to 1 as follows:

$$ConScore = 1 - \frac{x - min(x)}{max(x) - min(x)}$$

being x the addition of rank normalized values through evolution for each protein position. In such scale ConScore=0 represent the lowest value of conservation and ConScore=1 means that that protein position is fixed through the evolution.

Interface residues detection

Interface residues between subunits were determined using a distance threshold of 5Å in our in-silico models of the Respiratory complexes, using bio3d v-2.4-4 r-package.[42] All downstream analysis related to IFRs are referred to our in-silico models, that for CIV the alternative isoform composition is limited to COX4I1, COX6A1, COX6B1, COX7A2, COX7B, COX8A and NDUFA4, while the paralogous isoform included in the CV is ATP5MC1.

The detected IFRs were further analyzed to understand their evolutionary behavior. First, their conservation levels were compared to those of OxPhos residues not involved in binding sites, both across all subunit types and within individual subunit types, using the Wilcoxon test (for two distributions) or the Kruskal-Walli's test followed by Dunn's post hoc test. Second, conservation levels were assessed directly through ConScore values and indirectly by comparing AlphaMissense hotspot values within IFRs, depending on the subunit types involved in the junction.

ConScore benchmarking compared with AlphaMissense hotspot

We evaluate the performance of ConScore and AlphaMissense hotspot in identifying protein regions where mutations are more likely to be pathogenic. Their performance was measured as the area under Receiver Operating Characteristic (ROC) curve and the area under the Precision- Recall (PR) curve. For this purpose, we used described labeled missense mutations from ClinVar database that affect OxPhos genes. In this process of variant filter and consequence prediction bcftools version 1.20[48] and VEP. In this benchmarking variants considered pathogenic and likely pathogenic were both considered pathogenic, while benign and likely benign were considered as benign. For the area under the ROC and PR curve calculation and representation, we used ROCR version 1.0-11 and precrec version 0.14.4 R packages.[49,50]

Calculation of Probability of Structural Impact

To determine the degree of resilience of each protein position to mutations, it was performed a mutability scanner using membrane_predict_ddG.py tool from Pyrosetta 4.[43,51] This program mutates the wild-type residue of each position to the 19 alternative possibilities and calculates energy difference (ddG) for each mutation. As from a structural/functional point of view both stabilizing and destabilizing mutations may have an effect, we considered absolute values as a measure of the size effect. Finally, for each position we will have 19 absolute values of ddG and this measurement was performed for each protein position of all OxPhos components (not only mtDNA), according to our in-silico models, therefore regarding the presence of membrane environment, nuclear encoded neighboring subunits and prosthetic groups. Then, comparing residue specific 19 values distribution can be compared against the distribution of the same residue type (e.g. Alanine) from all subunits analyzed from our RC models. To compare both distributions, we used Kolmogorov-Smirnov non-parametric test, considering H_0 of being equal or lower absolute values of ddG, the probability of having a D statistic higher than expected is our PSI, for that position.

Assessment of the association between PSI and ConScore, human population constraints, RCs or subunit type

The study of the association between PSI and **ConScore levels** was performed firstly by Kruskal-Wallis test, followed by Dunn post hot test. For the assessment of association PSI and constraints among human populations, OxPhos positions were classified in two categories regarding whether amino-acid changes in such position were described or not as Fixed (SEPV=0) and Non-fixed positions (SEPV>0), where pair-wise comparison of PSI values between classes performing two-sample Wilcoxon tests. Structural resilience to amino-acid changes, measured as PSI, was also studied considering OxPhos RCs or subunit type (autosomal, mtDNA-encoded, or X-linked) using the Kruskal-Walli's test, followed by Dunn's post-hoc test. Dunn's post-hoc test was carried out using FSA R package version 0.9.6 (https://github.com/fishR-Core-Team/FSA), where FDR was corrected by Benjamini-Hochberg and a threshold of adjusted p-value$\leq$0.05 was set.

Evaluation of allelic imbalance (AIM) in OXPHOS genes at transcriptional level

To assess the existence of constraints for the expression of potential variability in nucleus-encoded OxPhos genes (nOXPHOS) determined by their diploid condition and heterozygosity, at transcriptional level, we analyzed single cell RNA-seq data (scRNAseq). To this end, it was developed an ad hoc methodology (Figure S7). First, fastq reads were mapped using STAR v-2-5-1[44] using reference GRCm38.91 for mouse samples. Then, duplicate reads in bam files were detected and tagged with markduplicate from Picard tools (http://broadinstitute.github.io/picard).

Next step mapped reads matching SNPs positions in OxPhos genes were retrieved for each cell. These SNPs information was obtained from the Mouse Genome Project.[38] Finally, a binomial test was carried out by cell and gene.

scRNAseq data

We performed the assessment of allele bias profiles from publicly available sources of scRNAseq data, from mouse samples (adult fibroblasts, embryonic neurons, and embryonic hepatocytes from different mouse F1 hybrids (males and females) from C57BL/6J x Castaneus crosses (GSE75659, GSE45719).[36,37] All these data were smartseq scRNAseq data. Regarding information needed to perform AIM analysis (samples' genotypes), SNPs used for allelic imbalance analysis of hybrids C57BL/6J x CAST/EiJ were acquired from the Mouse Genome Project.[38] Further details and characterization of these samples in Table S15.

AIM analysis

We designed an ad hoc strategy that needs double information from samples, genotype information to retrieve variants in heterozygosity and RNAseq data from single cell. Then for each cell, reads that matches SNPs positions were retrieved and annotated with harboring allele information, mapped gene, and cell id. From here, we used binomial test to assess allelic imbalance, considering the null hypothesis when both alleles have the same probability of expression (biallelic gene expression) and in consequence, the probability of expression of each allele is p=0.5 (Figure S9). The alternative hypothesis is that the alleles do not have the same probability of expression, the gene is monoallelic and hence the probability of expression of one of allele significantly higher than 0.5.

If we designate the number of reads of one specific allele as k and the total number of reads (considering both alleles) as n, then, under the null hypothesis, the number of reads for an allele should follow a binomial distribution with parameters n trials and probability of success of 0.5.

The p-value for the binomial test is calculated as the probability of observing a value as extreme or more extreme than the observed value of X, under the null hypothesis. It is calculated as:

$$P(X \geq k|H_0) = \sum_{i=k}^{n} C(n,k) \cdot 0.5^k \cdot (1 - 0.5)^{n-k}$$

False discovery rate was controlled adjusting p-values for multiple testing, by using the Benjamin-Hochberg method. Then only, results with an adjusted $p \leq 0.05$ as monoallelic expressed.

Furthermore, to evaluate AIM directed at a particular allele, which is a pattern observed in most cells at the gene level, we once again employed the binomial test. This test was applied only to genes that showed AIM results in at least 75% of the cells analyzed in the data sets under study (adult fibroblasts, E15 embryonic cells, and cells from early developmental stages). We tested three null hypotheses.

(1) H0: The count of cells with C57 AIM is equivalent to the count of cells without C57 AIM.
(2) H0: The count of cells with Cast AIM is equivalent to the count of cells without Cast AIM.
(3) H0: The count of biallelic cells is equivalent to the count of cells without biallelic cells.

In all these cases, under the null hypothesis (H0), the proportion of cells in the evaluated category is assumed to be p=0.5. If the observed proportion is significantly greater than 0.5, with a significance threshold of 0.05 ($p \leq 0.05$), we can assume the alternative hypothesis of enrichment in that category within our data set.

A caveat for this model is that relies over the assumption of single cell condition of the data, for biallelic outcomes. Nevertheless, the existence of doublets does not affect the confidence AI events toward a specific allele (which the hypothesis that we want to test).

QUANTIFICATION AND STATISTICAL ANALYSIS

Al statistical analysis or programmatic quantitative measurement was performed in Python o R and were described in their respective sections in method details or Figure legends.