An unusual amino acid substitution within hummingbird cytochrome *c* oxidase alters a key proton-conducting channel

Cory D. Dunn [a*], Bala Anı Akpınar [a], and Vivek Sharma [a,b*]

[a] Institute of Biotechnology
Helsinki Institute of Life Science
University of Helsinki
00014 Helsinki
Finland

[b] Department of Physics
University of Helsinki
00014 Helsinki
Finland

[*] Co-Corresponding Authorship

Cory Dunn, Ph.D.
P.O. Box 56
University of Helsinki
00014 Finland
Email: cory.dunn@helsinki.fi
Phone: +358 50 311 9307

Vivek Sharma, Ph.D.
P.O. Box 64
University of Helsinki
00014 Finland
Email: vivek.sharma@helsinki.fi
Phone: +358 50 5759 509

ABSTRACT

Hummingbirds in flight exhibit the highest metabolic rate of all vertebrates. The bioenergetic requirements associated with hovering flight raise the possibility of positive selection upon proteins encoded by hummingbird mitochondrial DNA. Here, we have identified a non-conservative change within the mitochondria-encoded cytochrome c oxidase subunit I (COI) that is fixed within hummingbirds, yet exceedingly rare among other metazoans. This unusual change can also be identified in several nectarivorous hovering insects, hinting at convergent evolution linked to diet or mode of flight over ~800 million years. We performed atomistic molecular dynamics simulations using bovine and hummingbird COI models, thereby bypassing experimental limitations imposed by the inability to modify mtDNA in a site-specific manner. Intriguingly, our findings suggest that COI amino acid position 153 provides control over the hydration and activity of a key proton channel. We discuss potential phenotypic outcomes for the hummingbird that are linked to this intriguing instance of positive selection upon the mitochondrial genome.

SIGNIFICANCE STATEMENT

How do organisms adapt to niches and environments that require unusual metabolic features? Changes to mitochondrial function are expected to be tightly linked to bioenergetic adaptation. Several proteins required for converting food into energy useful for the cell are specifically encoded by the mitochondrial genome, suggesting that adaptations required for exceptional metabolic performance might be found at this location. Here, we find that all hummingbirds harbor a remarkable change within their mitochondrial DNA that appears to be required for outstanding metabolic properties of this organism. Further analysis by computational simulations suggests that this hummingbird substitution alters proton movement across the mitochondrial inner membrane.

INTRODUCTION

Hummingbirds are distinguished by their use of hovering flight to feed upon nectar and insects, to defend their territories, and to carry out courtship displays (1–3). Their exceptional mobility demands a prodigious level of mitochondrial ATP synthesis, and indeed, the metabolic rate of hummingbird flight muscles is exceedingly high (4, 5). Many physiological and cellular features of hummingbirds appear to be tailored to their extreme metabolism, especially when considering that hummingbirds can be found within hypoxic environments up to 5000 meters above sea level (6). For example, hemoglobin structure (7) and cellular myoglobin concentration (8) appear to be adapted to the oxygen delivery needs of hummingbirds. Additionally, the hearts of hummingbirds are larger, relative to their body size, than other birds and can pump at a rate of more than 1000 beats per minute (9). Beyond ATP synthesis, the metabolism of these tiny endotherms must also buffer against heat loss

(4, 10, 11). At the subcellular level, adaptation to the need for increased ATP and heat production can be readily visualized, since mitochondria in hummingbird flight muscles are highly, perhaps maximally, packed with cristae and are found in close apposition to capillaries (12, 13). Hummingbirds have an exceptionally long lifespan when considering the allometric link between body mass and longevity (14), but whether hummingbird lifespan is linked to its unusual metabolic prowess is unclear.

Within the mitochondrial inner membrane, electrons progress through the electron transport chain (ETC), reach the cytochrome *c* oxidase (COX) complex, and are then used to reduce oxygen. Proton movements coupled to electron passage through COX contribute to the proton motive force (PMF) used for ATP production and thermogenesis (15, 16). While several COX subunits are nucleus-encoded and imported to mitochondria, the core, catalytic subunits of COX (subunits COI, COII, and COIII) are encoded by mitochondrial DNA (mtDNA) (17), raising the possibility that positive selection upon the mitochondrial genome may have contributed to the remarkable metabolic properties of hummingbirds. Here, we identify an amino acid substitution in COI that is universal among hummingbirds, yet exceedingly rare among other birds and vertebrates. Atomistic molecular dynamics (MD) simulations suggest that this substitution affects COX function and is likely to contribute to the uncommon physiological capabilities of hummingbirds.

RESULTS AND DISCUSSION

Hummingbird harbors unusual substitutions within the mitochondria-encoded subunit I of cytochrome c oxidase

We sought specific coding changes within mtDNA-encoded genes that might be associated with the extreme metabolic capabilities of hummingbirds. Toward this goal, we used software with the ability to predict ancestral sequences (18), along with our own custom software, to identify all amino acid positions mutated along the lineage leading to hummingbirds within a bird phylogenetic tree of concatenated, mtDNA-encoded protein sequences. Consistent with a need for mtDNA changes that permit the unusual metabolic properties of these animals, the lineage leading to the family Trochilidae exhibits the greatest number of changes when considering 635 internal edges (Figure 1A). Of those 208 positions altered along the edge leading to hummingbirds (Table S1), the most conserved amino acid position mutated during establishment of hummingbirds was COI position 153 (Figure 1B; for convenience, we use the amino acid numbering associated with the structurally characterized *Bos taurus* COI subunit). This non-conservative A153S substitution was universal among all 15 hummingbird COI sequences obtained from the RefSeq (19) database (Table S2), yet was absent from all other birds upon examination of an alignment of 645 Aves COI entries (Figure 1C).

Since COI is the most commonly used DNA sequence barcode for studying animal ecology and speciation (20, 21), we next analyzed additional sequences covering the COI region of interest that we obtained from the Barcode of Life Data (BOLD) server (22). Initially, we focused upon sequences from the bird order Apodiformes, a taxon encompassing hummingbirds and swifts. 914 of 915 informative samples annotated as hummingbird-derived were found to carry an A153S substitution at position 153 of COI (Table S3). The remaining sample is mis-annotated as hummingbird, as determined by BLASTN analysis of its barcode (23). In contrast, all 110 non-hummingbird Apodiformes samples harbored the ancestral A153. Extending our analysis to all informative bird barcodes, only 15/36,636 samples (< 0.1%) not annotated as hummingbird or its parental clade diverged from A at position 153. Assuming that these COI alterations were not the result of sequencing errors, we found that the identified changes to A153 outside of hummingbirds were not fixed within each identified genus (Table S4). No other COI change appears universally encoded by hummingbird mtDNA, and position 153 does not contact a nucleus-encoded subunit, suggesting the lack of a single compensatory change that would lead to substitution neutrality (24). Codons for alanine and serine are separated by a distance of only one base pair alteration, suggesting that sequence-level constraints do not explain the exceptional nature of the non-conservative A153S substitution in COI. Since A153 is nearly universal among birds, yet appears to be substituted for S in all hummingbirds, the A153S change within hummingbird COI is likely to be adaptive and to affect COX function.

Beyond birds, substitution for A at COI position 153 was also extremely unusual among chordates, a taxon encompassing vertebrates. Of 4,998 aligned Chordata sequences from the RefSeq dataset, only four non-hummingbird entries suggested a possible change at amino acid 153 (Table S5). Two RefSeq entries, from the sawtooth eel (*Serrivomer sector*) and the kuhli loach (*Pangio cf. anguillaris*), exhibit the A153S substitution characteristic of hummingbirds. However, further analysis of accumulated COI barcodes suggested that any substitution at position 153 is not widely shared among members of these vertebrate genera, in contrast to members of the hummingbird family, for which the A153S substitution appears universal. Extending our analysis to metazoans, substitution at A153 remains very rare. Indeed, only 146/7942 (< 2%) of informative RefSeq COI sequences harbor a substitution of A153 with any other amino acid (Table S6).

Evidence for convergent evolution toward a polar amino acid substitution at position 153 of cytochrome c oxidase subunit I

During our analysis of metazoan COI sequences, our attention was drawn to the prominent presence of A153S, and the similar non-conservative substitution A153T, in several bee species. Bees and hummingbirds are nectarivorous, thermogenic, and take advantage of energetically expensive hovering flight (1,

25). Moreover, bee metabolic rate relative to mass surpasses even that of hummingbird (26). Analysis of BOLD samples from hymenopteran families Apidae and Megachilidae, the "long-tongued" bees (27), indicate nearly 100% substitution at COI position 153 to either serine or threonine, while other families of bees harbor an ancestral alanine at position 153 (Table S7). Curiously, examination of COI sequences from millions of insect samples found in BOLD indicated that A153S and A153T conversion characterizes many, but not all genera within the Eristalini tribe of dipteran hoverfly (Table S8). Hoverflies within this clade very closely mimic bees visually and behaviorally (28). Together, our results hint at the exciting possibility of convergent evolution, potentially rooted in diet and foraging behavior, at the mitochondria-encoded COI.

Atomistic molecular dynamic simulations suggest that substitution at COI amino acid 153 has functional consequences for proton transport

In the high-resolution (1.5 Å) crystal structure (29) of COX from *Bos taurus* (> 87% identical in COI amino acid sequence to hummingbird), amino acid 153 is buried in the middle of COI transmembrane helix IV and is sandwiched between a water-filled cavity lined with polar residues and the D-channel, which is thought to conduct protons for oxygen reduction and PMF generation (30, 31) (Figure 2A). Moreover, residue 153 is only 10 Å from E242, a residue central to redox-coupled proton pumping by COX (15, 16). In an attempt to understand the functional relevance of the A153S change in hummingbirds, we performed atomistic classical MD simulations on two different vertebrate model systems at differing scales. Remarkably, multiple microsecond simulations demonstrated changes in hydration within the vicinity of position 153 that were coupled to the dynamics of the aforementioned E242. Specifically, during simulations of the entire 13-subunit wild-type bovine COX performed in membrane-solvent environment, E242 was typically found in the 'down' state ($\chi 2 \sim 60^{\circ}$), extending towards the D-channel proton uptake site D91 (Figures 2B and 2D). In contrast, upon A153S substitution, the bovine E242 commonly swung to the 'up' position ($\chi 2 \sim 180^{\circ}$, Figures 2C and 2D). Similar findings emerged (Figures 2E-G) from longer simulations performed on small bovine model systems, suggesting that the observed behavior is robust. The microscopic changes in hydration near E242 stabilized its 'up' position (Figures 2C and 2F) and resulted in its connection to the COI regions near the positively charged intermembrane space via water molecules (Figure S1). During simulations using a constructed hummingbird homology model, we saw that E242 behavior and channel hydration was dependent upon whether alanine or serine was present at position 153 (Figure 2H-I), although the effect was less prominent than in bovine models. In the constructed hummingbird model containing its wild-type S153 variant, E242 was stabilized in the 'down' position. Upon S153A replacement, both 'up' and 'down' populations were observed, and increased motility was visualized (Figure 2J) with corresponding changes to local hydration (Figure 2I).

Furthermore, our simulations suggest that the behavior of additional amino acids beyond E242 are affected by the amino acid found at position 153. For example, our hummingbird simulation strongly indicated that a change in F238 side chain angle is linked to E242 motion (Figure S2) and is influenced by whether residue 153 is an alanine or a serine. These data are supported by a GREMLIN co-evolution analysis (32), initiated by use of the bovine COI sequence, that suggested co-evolutionary coupling between F238 and E242 (Table S9). The behavior of another amino acid, F63, also appeared to depend upon the amino acid occupying position 153 (Figure S3A), stabilizing in the 'down' ($\chi1 \sim -160^{\circ}$) conformation in large bovine (Figure S3B), small bovine (Figure S3C), and small hummingbird (Figure S3D) A153 models. The 'up' conformation ($\chi1 \sim -77^{\circ}$) of F63 preferred in all models simulating S153 led to transient influx of water molecules in the domain above residue 153, in agreement with bovine COX structural data (Figures 2A and S3A). Taken together, MD analysis of naturally occurring variants, as performed here with a hummingbird COI substitution, clearly provides support for evolutionary and structural coupling between different regions of COI.

F238 has been suggested to play a key role in oxygen diffusion through COX (33), and indeed hummingbirds are characterized by a profound oxygen consumption rate during flight (4, 5). Therefore, altered behavior of F238 upon A153S substitution prompted us to consider the possibility that oxygen access to COX is augmented in the hummingbird. However, we are equivocal regarding a scenario in which improved oxygen access is prompted by A153S substitution. First, with caveats related to the evolutionary divergence between bacteria and vertebrates, a S153A substitution in bacterial COI (bacterial A-family cytochrome *c* oxidases harbor an S within their catalytic subunit) led to similar cytochrome *c* oxidation rates and initial proton pumping rates (34, 35). Moreover, the oxygen consumption rate of isolated hummingbird mitochondria, when normalized to mitochondrial inner membrane area, does not notably differ from mammalian mitochondria (12), suggesting similarity in the maximum rate of COX catalysis. Finally, despite any possible stretching of the oxygen channel linked to F238 movement, the access of oxygen to the active site is likely to be hampered by a corresponding 'up' flip of E242 and its surrounding hydration (Figure S2).

An additional COI variant characterizes, but is not ubiquitous among, hummingbirds

Within the class Aves, one other COI variant beyond S153 appears restricted to hummingbirds (Tables S1). Among the 15 hummingbird COI sequences found within the RefSeq database, nine contained a conservative V83I substitution that is found in no other bird entry (Table S2). Expanding our analysis to Apodiformes barcodes obtained from BOLD, 110/110 non-hummingbird samples carried the V83 allele ancestral for birds. In contrast, 671/929 informative hummingbird samples within this dataset carried a V83I substitution, and 258/929 samples harbored a valine at this

position (Table S3). Looking more broadly at 36,878 bird samples, substitution at V83 is extremely rare among birds (< 0.1%), although unlike the A153S substitution, this V83I allele may be widely shared among members of a very limited number of non-hummingbird genera (Table S4). Phylogenetic analysis suggests that the V83I substitution was present at the establishment of hummingbirds, then subsequently reverted to valine several times during expansion of this clade (Figure S4). An analysis of BOLD entries performed blind to hummingbird species and details of sample acquisition suggests that hummingbirds encoding V83 or I83 can be found within overlapping geographical regions, yet there is a trend toward recovery of the I83 allele at higher altitudes (Figure S5 and Table S10). Within the COX enzyme, amino acid 83 lies within 9 Å of COI residue D91, which contributes to proton uptake via the eponymous D-channel described above (Figure 2A). Position 83 is also located within 6 Å of N80, a component of the 'asparagine gate' which is thought to move protons toward the active site (36–38).

What is the phenotypic outcome of the A153S substitution in COI?

Our results show clear changes to the behavior of key D-channel residues and to surrounding hydration when comparing different COI models harboring alanine or serine at position 153, suggesting altered proton movement at hummingbird COX. Interestingly, previous studies of bacterial respiration suggest that amino acid 153 can influence coupling between electron transport and COX proton pumping, further indicating that proton motility is a focus of selection during evolution of hummingbirds. Specifically, a serine to aspartic acid change made at the analogous amino acid position in *Rhodobacter sphaeroides* COI abolished proton pumping while allowing electron transport coupled to proton transfer from the periplasmic side of the bacterial inner membrane when the enzyme was analyzed under zero PMF conditions (34, 39). Further suggesting strong selection on proton handling by COX, the hummingbird-specific V83I substitution is located near the 'asparagine gate' at the matrix side of the mitochondrial inner membrane, and mutations near this site lead to changes in the number of protons pumped per oxygen reduction (35). Also of note, functional links have been suggested to exist (40) between the asparagine gate and the key E242 residue, the behavior of which is clearly affected by A153S mutation.

We suggest two potential outcomes of the D-channel changes prompted by the A153S change universal to hummingbirds. First, if the bovine models accurately reflect the outcome of this substitution, hydration differences associated with the presence of a polar residue at position 153 may promote intrinsic uncoupling (41) of COX when the PMF is high across the mitochondrial inner membrane, even leading to the use of protons from the intermembrane space for oxygen reduction under conditions of high polarization (42). Rapid, on-site decoupling of proton pumping from electron transport may serve as a local response to cessation of flight,

when an immediate rise in matrix ATP might lead to excessive PMF, ETC over-reduction, and reactive oxygen species (ROS) production (43, 44). Intrinsic, local, and immediate uncoupling might be particularly important within hummingbird muscle, where the densely packed ETC components might generate a destructive wave of ROS linked to inner membrane hyperpolarization.

Furthermore, hummingbirds require robust thermoregulation, since these organisms have high surface to mass ratios, can be found at elevations associated with low temperatures, and are subject to convective heat loss while engaged in hovering flight (3, 6, 45). Beyond the possibility of heat generation by COX 'slip', or decoupling of electron transport from proton pumping (46), results emerging from our bovine models raise the possibility that changes to COI hydration accompanying A153S substitution might allow direct, albeit limited movement of protons across the inner membrane that could contribute to non-shivering thermogenesis under conditions of high PMF. Our findings regarding the altitudes at which COI variants at position 83 can be found are also consistent with a role for the D-channel in hummingbird heat generation. Interestingly, thermoregulation may act as an initial selective force toward increased metabolic capacity (47) and therefore may have played a particularly prominent role during the evolution of hummingbirds.

Thus far, the vertebrate mitochondrial genome remains refractory to directed modification toward a desired sequence change (48), preventing a direct test of the hummingbird-enriched COI substitutions in the background of a hummingbird mtDNA or of a related, non-hummingbird mitochondrial genome. However, future biochemical experiments guided by our combined use of phylogenetic analysis and atomistic simulations may be informative regarding the role of hummingbird COI changes that have emerged from positive selection, once thought unlikely to shape mitochondria-encoded genes (49). Excitingly, other changes to mtDNA-encoded oxidative phosphorylation machinery beyond COX are rare among birds yet common in hummingbirds, and these substitutions await further analysis. Finally, while mtDNA sequence is far more prevalent, we expect that accumulating nuclear DNA sequence information from birds (50) will allow analysis of divergent nucleus-encoded members of the oxidative phosphorylation machinery and their potential role in hummingbird metabolism.

METHODOLOGY

Sequence acquisition, alignment, phylogenetic analysis, and annotation

Mitochondrial proteomes were downloaded from the NCBI RefSeq database (release 92) (19). Taxonomy analysis was performed using the 'taxize' package (51) and the NCBI Taxonomy database (52), with manual curation when required. Beyond COI sequences acquired from the RefSeq database, additional COI barcodes were retrieved from the BOLD server (22).

Alignments were performed by use of standalone MAFFT (version 7.407) (53) or by T-coffee (version 13.40.5) in regressive mode (54). For initial alignments of insect COI barcodes, MAFFT alignment was performed using an online server (55, 56), and translations of barcodes using the appropriate codon table were performed using AliView (57).

For analysis of mutations along tree edges, a maximum-likelihood phylogenetic tree based upon an alignment of concatenated sequences of mtDNA-encoded proteins from the RefSeq database was generated by FastTreeMP (version 2.1.11) (58), then FigTree (version 1.4.4, https://github.com/rambaut/figtree/releases) was used to root the resulting tree on the edge between birds and *Bos taurus*. The alignment and rooted tree were then used as input by PAGAN (version 0.61) (18) for the purpose of ancestral reconstruction. The "binary-table-by-edges-v2.1py" script was used to generate a table reporting upon whether a given position was mutated on each tree edge, and the "add-convention-to-binarytable-v.1.1.py" script was used to apply *Bos taurus* position information to the output. The predicted ancestral and descendant values at the edge leading to hummingbirds were generated using the script "report-on-F-values-v1.1.py", and all possible characters that could be found at each amino acid position of interest across the entire tree (including *Bos taurus*) were extracted by the script "extract-position-values_species_and_nodes-v1.1.py". Scripts developed and used during this study can be found at https://github.com/corydunnlab/hummingbird.

Elevation data matching geolocation latitude and longitude were acquired from the Gpsvisualizer site (www.gpsvisualizer.com).

Modeling and simulation

We constructed small and large model systems of bovine mitochondrial cytochrome *c* oxidase from the high-resolution (1.5 Å) crystal structure (PDB 5B1A) (29). The large model system comprised all thirteen subunits, whereas in the small model only two catalytic subunits (COI and COII) were included, thereby allowing longer timescale simulations. Both wild-type and mutant (A153S) cases were considered. The larger protein system was embedded in a multicomponent lipid bilayer (POPC:POPE:cardiolipin in 5:3:2 ratio) and only single component bilayer (POPC) was used in the case of two subunit complex, both constructed using CHARMM-GUI (59). Solvent water and Na^+ and Cl^- ions (150 mM each) were added. In both setups, metal centers were in oxidized states with a water and an hydroxyl ligand at heme a_3 and Cu_B, respectively. Crosslinked Y244 was anionic [see also (60)]. All amino acids were in their standard protonation states, except E242, K319 and D364 of subunit I, which were kept in a neutral state. The CHARMM force field parameters for protein, water, lipids and metal centers were taken from (61–63). Additional subunit I/II homology models of hummingbird cytochrome *c* oxidase [both wild-type (S153) and mutant (A153)

model systems] were constructed using bovine structure (PDB: 5B1A) and the predicted amino acid sequence of *Florisuga mellivora* (accession YP_009155396.1).

All MD simulations were performed with GROMACS software (64). Prior to production runs, all model systems were subjected to energy minimization followed by an equilibration MD. During equilibration, Nose-Hoover thermostat (65, 66) and Berendsen barostat (67) were applied. LINCS (68) algorithm implemented in GROMACS was applied to achieve the time step of 2 fs. During production runs Nose-Hoover thermostat and Parinello-Rahman barostat (69) were applied to keep temperature and pressure constant at 310 K and 1 atm, respectively. The large and small model systems of bovine oxidase were simulated for 1.5 and 3 μs, respectively. The hummingbird COX wild-type and mutant models were simulated for 1 μs each, resulting in a total of 11 μs of atomistic simulations. VMD (70) was used for the visualization of trajectories and analysis.

DATA AVAILABILITY

All data related to this manuscript are available upon request.

ACKNOWLEDGEMENTS

C.D.D. is funded by an ERC Starter Grant (RevMito 637649) and by the Sigrid Jusélius Foundation. V.S. is funded by the Academy of Finland, the University of Helsinki, and the Sigrid Jusélius Foundation. We thank Fyodor Kondrashov for advice on COI sequence recovery and alignment, Gregor Habeck for assistance in use of R Studio, and Aapo Malkamäki for support in MD simulation setup. We also acknowledge the Center for Scientific Computing, Finland for their generous computational support.

AUTHOR CONTRIBUTIONS

C.D.D. and B.A.A. performed phylogenetic and taxonomic analyses. V.S. performed molecular dynamics simulations. All authors prepared the manuscript text and figures.

COMPETING INTERESTS

The authors declare no competing interests.

FIGURE LEGENDS

Figure 1: A rare alanine to serine substitution at bovine COI position 153 is universal among hummingbirds. (A) The edge leading to hummingbird exhibits the largest number of substitutions within mitochondria-encoded proteins among all internal edges in a bird phylogenetic tree. A maximum likelihood tree was generated from an alignment of concatenated mitochondrial proteins from birds and *Bos taurus* using T-coffee in regressive mode (54), followed by ancestral prediction using PAGAN (18). Amino acid substitutions between each pair of ancestral and descendant nodes internal to the bird tree (node-to-node) were determined, summed across all positions, and plotted. (B) Among those changes found within

the edge leading to hummingbirds, substitution at COI position 153 is most infrequent among birds, occurring only once. A plot demonstrating the number of times a given amino acid position was altered within the bird phylogeny is provided. (C) Serine at COI position 153 is unique to, and universal among, hummingbirds, as confirmed by phylogenetic analysis and by examination of an alignment of 645 Aves COI entries. Bird orders are arranged based upon a supertree modified from (71) under a Creative Commons license.

Figure 2: Hydration-coupled dynamics of conserved residue E242 are altered by the A153S substitution found in hummingbird COI. (A) The D-channel of proton transfer is located near residue 153 (dotted circles) in the high-resolution crystal structure of COX from *Bos taurus* (PDB 5B1A). Crystallographically resolved water molecules (purple spheres) in the domain above A153, together with nearby polar amino acids, form a potential proton transfer path. Cu_B is shown in orange and high spin heme in yellow. The catalytic COI subunit is shown with transparent ribbons and amino acids are displayed in atom-based coloring. (B-D) COI hydration and E242 side chain position are altered by substitution at COI position 153 in a large bovine COX simulation. (B) illustrates the native structure (A153) and (C) demonstrates the effects of A153S substitution. A red arrow highlights major changes to hydration, and water occupancy is shown as an orange colored mesh at an isovalue of 0.15. (D) E242 side chain dihedral angle (χ2) within COI encoding A153 (black) or S153 (red) during 1.5 μs of bovine large model simulation is displayed. Here, E242 adopts a predominant 'up' conformation within A153S substituted COI. (E-G), as in (B-D), but a small bovine model simulation has been deployed. (H-J), as in (B-D), but a small hummingbird model has been simulated. In (J), S153 (red) is wild-type and A153 (black) is mutant.

SUPPLEMENTARY INFORMATION

Figure S1: Water-based paths toward the positively charged intermembrane space. Due to shift in 'down' to 'up' conformation of E242, the side chain of this amino acid connects to the hydrophilic region above heme propionates and near amino acids Y231 and T146 following A153S substitution in a small bovine COI simulation.

Figure S2: The position of the COI F238 side chain is coupled to E242 dynamics. (A) F238 is shown within the hummingbird small model simulation when E242 is in the (A) down position and when E242 points in the (B) up position. F238 dihedral angle (χ1, N-CA-CB-CG) is compared to E242 dihedral angle (χ2, CA-CB-CG-CD) in the (C) bovine big model simulation and in the (D) hummingbird small model simulation. E242 data from Figures 2D and 2J are shown, with F238 dihedral angle for the A153 variant plotted in blue and for the S153 variant plotted in orange.

Figure S3: The dynamic behavior of the conserved F63 is determined by amino acid identity at position 153. (A) Simulation

snapshots showing two positions of the F63 sidechain within the bovine COX are shown in green licorice. The 'up' position (χ1 ~ -77°) results in local hydration in the case of A153S substitution mutant, whereas the stable 'down' conformation (χ1 ~ -160°) that blocks the entry of water molecules, is stabilized in wild-type case. The preferred angle of F63 during MD simulations is altered when alanine is swapped for serine in large (B) and small (C) model system of bovine COX. F63 dynamics are similarly altered in a small model system of hummingbird oxidase (D). A153 (black) and S153 (red).

Figure S4: COI I83 is ancestral for hummingbirds and reverts to the bird consensus V83 in multiple lineages. Reference mtDNA sequences (release 91) of the indicated organisms were obtained and aligned by MAFFT using the G-INS-i iterative refinement method. Next, a phylogenetic analysis was performed within the MEGA7 software (72) using a maximum likelihood approach based upon the Tamura-Nei model (73). The amino acid at COI position 83 is illustrated next to each organism within the resulting tree.

Figure S5: A meta-analysis suggests that COI I83 may be associated with hummingbird habitation at higher elevation. Violin plots were generated using elevation data listed in Table S10. Median is denoted by the dotted line, with quartiles illustrated by solid black lines.

Table S1: Analysis of the frequency at which mutations occur at specific positions mutated along the lineage to hummingbirds. Following concatenation and alignment of protein sequences encoded by birds and by *Bos taurus*, a maximum-likelihood tree was generated, and 208 positions harboring substitutions that occurred along the lineage to hummingbirds were predicted. The total number of times each position was found to mutate within birds is provided, along with the predicted ancestral and descendant amino acids on the edge leading to hummingbirds, and all amino acid possibilities at that position (including within *Bos taurus*). Whether substitutions at a given position are detected at edges either within or leading to the hummingbird clade is also provided.

Table S2: Analysis of full-length bird COI sequences obtained from the RefSeq database. Sequences of all mitochondria-encoded proteins found in the NCBI RefSeq database (release 92) (19) were downloaded, and the COI FASTA sequences of 645 birds were extracted. MAFFT alignment was performed (53) using the G-INS-i iterative refinement approach, and those birds with a A153S substitution are listed along with the variant harbored by each species at COI position 83.

Table S3: Examination of amino acids 83 and 153 in Apodiformes barcodes obtained from BOLD. The query "Apodiformes" was made using the BOLD server (22) to recover COI FASTA sequences. MAFFT alignment was carried out using L-INS-i iterative refinement method and translated in AliView (57) using the vertebrate mtDNA translation table.

Informative samples with annotated species names were assessed to determine the amino acids found at positions 83 and 153.

Table S4: Analysis of Aves COI barcodes obtained from BOLD. The query "Aves" was made using the BOLD server (22) to recover COI FASTA sequences. MAFFT alignment was carried out under the "auto" setting and translated in AliView (57) using the vertebrate mtDNA translation table. Informative COI samples outside of the order Apodiformes and not harboring alanine at position 153 or valine at position 83 are listed.

Table S5: Analysis of chordate COI sequences obtained from the RefSeq database. COI sequences obtained from the RefSeq database (19) were filtered for chordate entries. MAFFT alignment was performed using the "auto" setting and poorly aligned sequences were deleted. Hummingbird sequences were removed. Four informative, non-hummingbird sequences were characterized by substitution of A153 for other amino acids, although analysis of annotated BOLD accessions labelled with a species name or entry suggest that these COI substitutions are not generally shared throughout each genus.

Table S6: Analysis of metazoan COI sequences obtained from the RefSeq database. COI sequences obtained from the RefSeq database (19) were filtered for metazoan entries. Analysis was performed as in Table S5, and non-chordate samples with a substitution at A153 are displayed.

Table S7: Analysis of bee COI barcodes obtained from BOLD. Entries for bee families Apidae, Megachilidae, Colletidae, Halictidae, Andrenidae, and Melittidae, as defined in (74), were obtained from BOLD (22), but no samples from the small bee family Stenotritidae were available. In addition, no Spheciformes samples were found in BOLD. After deletion of samples not of the COI-5P class or not associated with a species, MAFFT alignment was carried out under the "auto" setting. DNA sequences were translated to the informative reading frame in AliView (57) using the invertebrate mtDNA translation table. Substitution quantification for each family is provided, along with the variant found at position 153 in each sample.

Table S8: Analysis of COI barcodes from Eristalini tribe of hoverflies obtained from BOLD. Entries for the Anasimyia, Eristalinus, Eristalis, Helophilus, Lejops, Mesembrius, Palpada, Parhelophilus, Phytomia, Senaspis, Mallota, Chasmomma genera were obtained and analyzed as in Table S7.

Table S9: GREMLIN analysis of co-evolution. GREMLIN (32) was run using the default settings and the bovine COI sequence as a query. 'i_id' and 'j_id' represent analyzed residues, 'r-sco' represents the raw scoring, 's-sco' represents the scaled score generated from the raw score and average of the raw score, and 'prob' represents the probability of residue contact given the scaled score and the sequences per length.

Table S10: Geolocation data for COI variants found at position 83. For each BOLD accession associated with a hummingbird species, an elevation was derived from Gpsvisualizer. Elevations provided by BOLD (22) were not utilized, since they typically did not match with the BOLD-annotated longitude and latitude and were sparse within the dataset.

REFERENCES

1. Norberg UM (1996) Energetics of flight. *Avian Energetics and Nutritional Ecology*, ed Carey C (Chapman & Hall, New York), pp 199–249.
2. Hainsworth FR, Wolf LL (1972) Power for Hovering Flight in Relation to Body Size in Hummingbirds. *Am Nat* 106(951):589–596.
3. Altshuler DL, Dudley R (2002) The ecological and evolutionary interface of hummingbird flight physiology. *J Exp Biol* 205(Pt 16):2325–36.
4. Lasiewski RC (1963) Oxygen consumption of torpid, resting, active, and flying hummingbirds. *Physiol Zool* 36(2):122–140.
5. Suarez RK, Brown GS, Hochachka PW (1986) Metabolic sources of energy for hummingbird flight. *Am J Physiol* 251(3 Pt 2):R537-42.
6. Projecto-Garcia J, et al. (2013) Repeated elevational transitions in hemoglobin function during the evolution of Andean hummingbirds. *Proc Natl Acad Sci U S A* 110(51):20669–74.
7. Projecto-Garcia J, et al. (2013) Repeated elevational transitions in hemoglobin function during the evolution of Andean hummingbirds. *Proc Natl Acad Sci* 110(51):20669–20674.
8. Johansen K, Berger M, Bicudo JEPW, Ruschi A, de Almeida PJ (1987) Respiratory Properties of Blood and Myoglobin in Hummingbirds. *Physiol Zool* 60(2):269–278.
9. Bishop CM (1997) Heart mass and the maximum cardiac output of birds and mammals: Implications for estimating the maximum aerobic power input of flying animals. *Philos Trans R Soc B Biol Sci* 352(1352):447–456.
10. Suarez RK, Gass CL (2002) Hummingbird foraging and the relation between bioenergetics and behaviour. *Comp Biochem Physiol Part A Mol Integr Physiol* 133(2):335–343.
11. López-Calleja MV, Bozinovic F (1995) Maximum Metabolic Rate, Thermal Insulation and Aerobic Scope in a Small-Sized Chilean Hummingbird (Sephanoides sephanoides). *Auk* 112(4):1034–1036.
12. Suarez RK, Lighton JR, Brown GS, Mathieu-Costello O (1991) Mitochondrial respiration in hummingbird flight muscles. *Proc Natl Acad Sci U S A* 88(11):4870–3.
13. Mathieu-Costello O, Suarez RK, Hochachka PW (1992) Capillary-to-fiber geometry and mitochondrial density in hummingbird flight muscle. *Respir Physiol* 89(1):113–32.
14. Calder III WA (1990) Avian Longevity

and Aging. *Genetic Effects on Aging*, ed Harrison DE (Telford Press, West Caldwell, NJ, USA), pp 185–204.

15. Wikström M, Sharma V (2018) Proton pumping by cytochrome c oxidase - A 40 year anniversary. *Biochim Biophys acta Bioenerg* 1859(9):692–698.
16. Wikström M, Krab K, Sharma V (2018) Oxygen Activation and Energy Conservation by Cytochrome c Oxidase. *Chem Rev* 118(5):2469–2490.
17. Johnston IG, Williams BP (2016) Evolutionary inference across eukaryotes identifies specific pressures favoring mitochondrial gene retention. *Cell Syst* 2(2):101–111.
18. Löytynoja A, Vilella AJ, Goldman N (2012) Accurate extension of multiple sequence alignments using a phylogeny-aware graph algorithm. *Bioinformatics* 28(13):1684–1691.
19. O'Leary NA, et al. (2016) Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* 44(D1):D733–D745.
20. Pentinsaari M, Salmela H, Mutanen M, Roslin T (2016) Molecular evolution of a widely-adopted taxonomic marker (COI) across the animal tree of life. *Sci Rep* 6(October):1–12.
21. Hill GE (2015) Mitonuclear Ecology. *Mol Biol Evol* 32(8):1917–27.
22. Ratnasingham S, Hebert PDN (2007) bold: The Barcode of Life Data System (http://www.barcodinglife.org). *Mol Ecol Notes* 7(3):355–364.
23. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–10.
24. Blier PU, Dufresne F, Burton RS (2001) Natural selection and the evolution of mtDNA-encoded peptides: evidence for intergenomic co-adaptation. *Trends Genet* 17(7):400–6.
25. Loli D, Bicudo JEPW (2005) Control and regulatory mechanisms associated with thermogenesis in flying insects and birds. *Biosci Rep* 25(3–4):149–80.
26. Suarez RK, Staples JF, Lighton JR, Mathieu-Costello O (2000) Mitochondrial function in flying honeybees (Apis mellifera): respiratory chain enzymes and electron flow from complex III to oxygen. *J Exp Biol* 203(Pt 5):905–11.
27. Cariveau DP, et al. (2016) The Allometry of Bee Proboscis Length and Its Uses in Ecology. *PLoS One* 11(3):e0151482.
28. Golding YC, Edmunds M (2000) Behavioural mimicry of honeybees (Apis mellifera) by droneflies (Diptera: Syrphidae: Eristalis spp.). *Proc R Soc London Ser B Biol Sci* 267(1446):903–909.
29. Yano N, et al. (2016) The Mg 2+ -containing Water Cluster of Mammalian Cytochrome c Oxidase Collects Four Pumping Proton Equivalents in Each Catalytic Cycle. *J Biol Chem* 291(46):23882–23894.
30. Kaila VRI, Sharma V, Wikström M (2011) The identity of the transient proton loading site of the proton-pumping mechanism of cytochrome c oxidase.

Biochim Biophys Acta - Bioenerg 1807(1):80–84.

31. Wikström M, Verkhovsky MI (2011) The D-channel of cytochrome oxidase: An alternative view. *Biochim Biophys Acta - Bioenerg* 1807(10):1273–1278.
32. Kamisetty H, Ovchinnikov S, Baker D (2013) Assessing the utility of coevolution-based residue-residue contact predictions in a sequence- and structure-rich era. *Proc Natl Acad Sci* 110(39):15674–15679.
33. Mahinthichaichan P, Gennis RB, Tajkhorshid E (2018) Cytochrome aa3 Oxygen Reductase Utilizes the Tunnel Observed in the Crystal Structures To Deliver O2 for Catalysis. *Biochemistry* 57(14):2150–2161.
34. Namslauer A, et al. (2007) Plasticity of Proton Pathway Structure and Water Coordination in Cytochrome c Oxidase. *J Biol Chem* 282(20):15148–15158.
35. Pfitzner U, et al. (2000) Tracing the D-Pathway in Reconstituted Site-Directed Mutants of Cytochrome c Oxidase from Paracoccus denitrificans †. *Biochemistry* 39(23):6756–6762.
36. Henry RM, Yu C-H, Rodinger T, Pomès R (2009) Functional Hydration and Conformational Gating of Proton Uptake in Cytochrome c Oxidase. *J Mol Biol* 387(5):1165–1185.
37. Liang R, Swanson JMJ, Wikström M, Voth GA (2017) Understanding the essential proton-pumping kinetic gates and decoupling mutations in cytochrome c oxidase. *Proc Natl Acad Sci* 114(23):5924–5929.
38. Ghane T, et al. (2018) Hydrogen-Bonded Network and Water Dynamics in the D-channel of Cytochrome c Oxidase. *J Membr Biol* 251(3):299–314.
39. Pawate AS, et al. (2002) A mutation in subunit I of cytochrome oxidase from Rhodobacter sphaeroides results in an increase in steady-state activity but completely eliminates proton pumping. *Biochemistry* 41(45):13417–23.
40. Vakkasoglu AS, Morgan JE, Han D, Pawate AS, Gennis RB (2006) Mutations which decouple the proton pump of the cytochrome c oxidase from Rhodobacter sphaeroides perturb the environment of glutamate 286. *FEBS Lett* 580(19):4613–4617.
41. Murphy MP, Brand MD (1988) The stoichiometry of charge translocation by cytochrome oxidase and the cytochrome bc1 complex of mitochondria at high membrane potential. *Eur J Biochem* 173(3):645–51.
42. Mills DA, Schmidt B, Hiser C, Westley E, Ferguson-Miller S (2002) Membrane potential-controlled inhibition of cytochrome c oxidase by zinc. *J Biol Chem* 277(17):14894–901.
43. Papa S, Guerrieri F, Capitanio N (1997) A possible role of slips in cytochrome C oxidase in the antioxygen defense system of the cell. *Biosci Rep* 17(1):23–31.
44. Kadenbach B (2003) Intrinsic and extrinsic uncoupling of oxidative phosphorylation. *Biochim Biophys Acta* 1604(2):77–94.

45. Altshuler DL, Dudley R, McGuire JA (2004) Resolution of a paradox: hummingbird flight at high elevation does not come without a cost. *Proc Natl Acad Sci U S A* 101(51):17731–6.
46. Musser SM, Chan SI (1995) Understanding the cytochrome c oxidase proton pump: thermodynamics of redox linkage. *Biophys J* 68(6):2543–2555.
47. Bennett AF (1991) The evolution of activity capacity. *J Exp Biol* 160:1–23.
48. Patananan AN, Wu TH, Chiou PY, Teitell MA (2016) Modifying the Mitochondrial Genome. *Cell Metab* 23(5):785–796.
49. Little AG, Lougheed SC, Moyes CD (2012) Evolution of mitochondrial-encoded cytochrome oxidase subunits in endothermic fish: The importance of taxon-sampling in codon-based models. *Mol Phylogenet Evol* 63(3):679–684.
50. Zhang G (2015) Bird sequencing project takes off. *Nature* 522(7554):34.
51. Chamberlain SA, Szöcs E (2013) taxize: taxonomic search and retrieval in R. *F1000Research* 2:191.
52. Federhen S (2012) The NCBI Taxonomy database. *Nucleic Acids Res* 40(D1):D136–D143.
53. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30(4):772–80.
54. Garriga E, et al. (2018) Fast and accurate large multiple sequence alignments using root-to-leave regressive computation. *bioRxiv*:490235.
55. Katoh K, Rozewicki J, Yamada KD (2017) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* (June):1–7.
56. Kuraku S, Zmasek CM, Nishimura O, Katoh K (2013) aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. *Nucleic Acids Res* 41(W1):W22–W28.
57. Larsson A (2014) AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* 30(22):3276–8.
58. Price MN, Dehal PS, Arkin AP (2010) FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS One* 5(3):e9490.
59. Lee J, et al. (2016) CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J Chem Theory Comput* 12(1):405–413.
60. Malkamäki A, Sharma V (2019) Atomistic insights into cardiolipin binding sites of cytochrome c oxidase. *Biochim Biophys Acta - Bioenerg* 1860(3):224–232.
61. Klauda JB, et al. (2010) Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J Phys Chem B* 114(23):7830–43.
62. Best RB, et al. (2012) Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ(1)

and χ(2) dihedral angles. *J Chem Theory Comput* 8(9):3257–3273.

63. MacKerell AD, et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102(18):3586–616.
64. Abraham MJ, et al. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1–2:19–25.
65. Nosé S (1984) A unified formulation of the constant temperature molecular dynamics methods. *J Chem Phys* 81(1):511–519.
66. Hoover WG (1985) Canonical dynamics: Equilibrium phase-space distributions. *Phys Rev A* 31(3):1695–1697.
67. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. *J Chem Phys* 81(8):3684–3690.
68. Hess B (2008) P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J Chem Theory Comput* 4(1):116–122.
69. Parrinello M, Rahman A (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *J Appl Phys* 52(12):7182–7190.
70. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. *J Mol Graph* 14(1):33–38.
71. Davis KE, Page RDM (2014) Reweaving the tapestry: a supertree of birds. *PLoS Curr* 6(506):1–91.
72. Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 33(7):1870–4.
73. Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 10(3):512–26.
74. Sann M, et al. (2018) Phylogenomic analysis of Apoidea sheds new light on the sister group of bees. *BMC Evol Biol* 18(1):71.

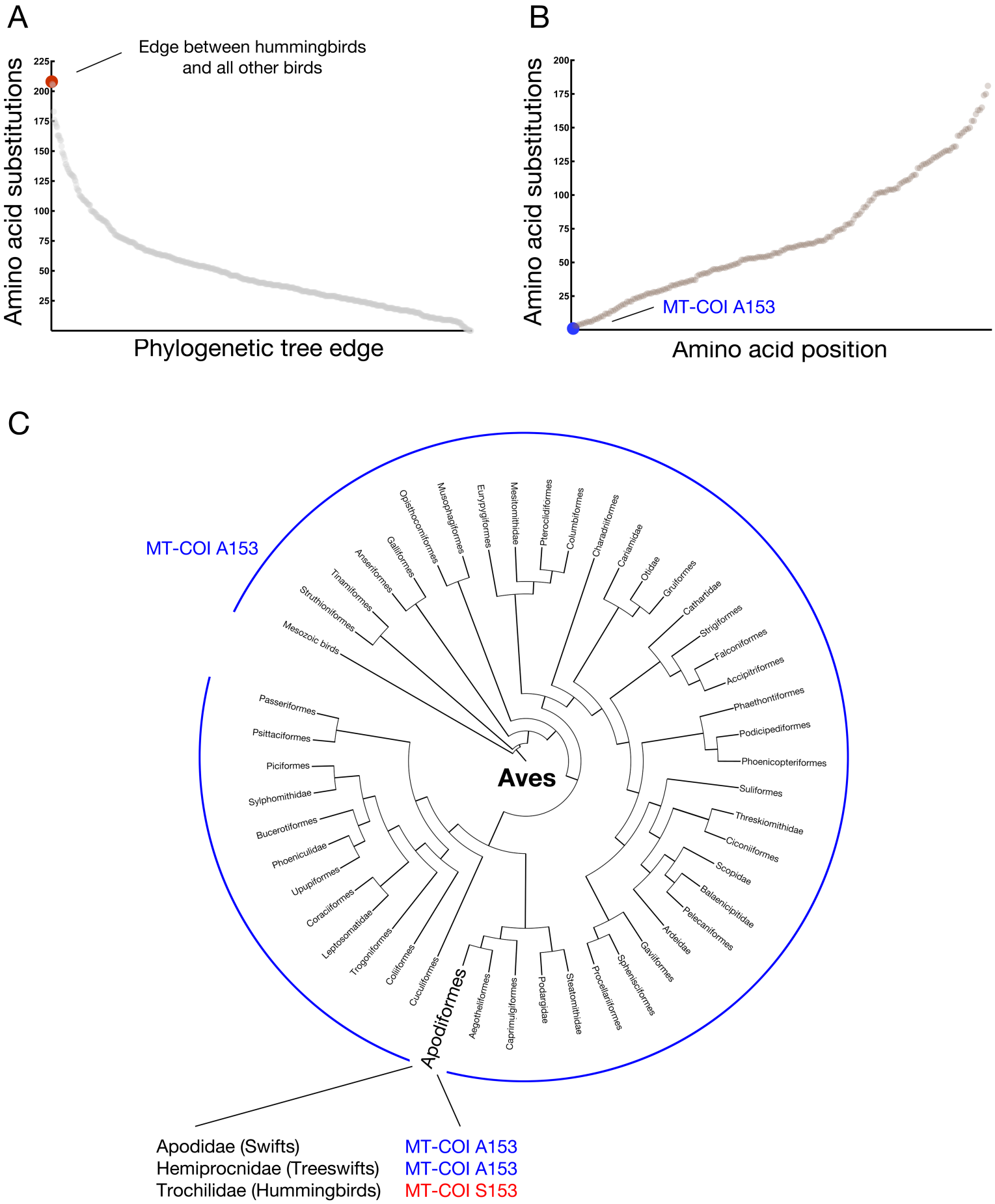
A
Edge between hummingbirds and all other birds
Amino acid substitutions
225
200
175
150
125
100
75
50
25
Phylogenetic tree edge
B
Amino acid substitutions
200
175
150
125
100
75
50
25
MT-COI A153
Amino acid position
C
MT-COI A153
Aves
Apodiformes
Opisthocomiformes
Musophagiformes
Eurypygiformes
Mesitornithidae
Pteroclidiformes
Columbiformes
Charadriiformes
Cariamidae
Otidae
Gruiformes
Cathartidae
Strigiformes
Falconiformes
Accipitriformes
Phaethontiformes
Podicipediformes
Phoenicopteriformes
Suliformes
Threskiornithidae
Ciconiiformes
Scopidae
Balaenicipitidae
Pelecaniformes
Ardeidae
Gaviiformes
Sphenisciformes
Procellariformes
Steatornithidae
Podargidae
Caprimulgiformes
Aegotheliformes
Cuculiformes
Coliiformes
Trogoniformes
Leptosomatidae
Coraciiformes
Upupiformes
Phoeniculidae
Bucerotiformes
Sylphomithidae
Piciformes
Psittaciformes
Passeriformes
Mesozoic birds
Struthioniformes
Tinamiformes
Anseriformes
Galliformes
Apodidae (Swifts) MT-COI A153
Hemiprocnidae (Treeswifts) MT-COI A153
Trochilidae (Hummingbirds) MT-COI S153

Figure 1

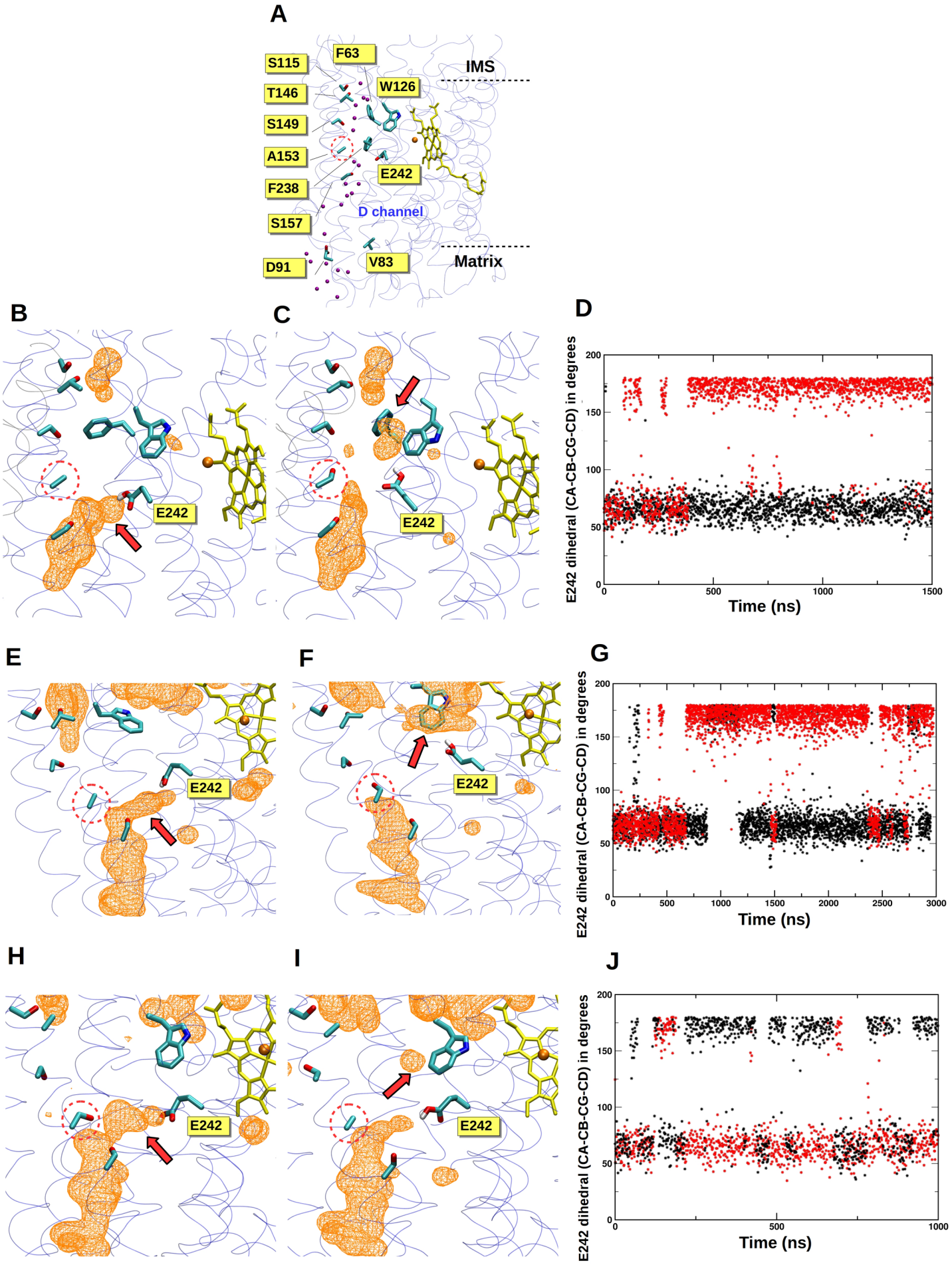
A
S115
F63
IMS
T146
W126
S149
A153
E242
F238
D channel
S157
D91
V83
Matrix
B
C
D
E242
E242
E242 dihedral (CA-CB-CG-CD) in degrees
Time (ns)
E
F
G
E242
E242
E242 dihedral (CA-CB-CG-CD) in degrees
Time (ns)
H
I
J
E242
E242
E242 dihedral (CA-CB-CG-CD) in degrees
Time (ns)

Figure 2