

Hydrophobic and aerophobic mitochondrial mutational spectrum of birds

Gusarov Y.^{1*}, Mikhailova A.¹, Efimenko B.¹, Gunbin K.¹, Bushuev A.², Burskaya V.³, Popadin K.¹

¹ Center of the Genomic Researches, Higher School of Living Systems, Institute of Medicine and Life Science, Immanuel Kant Baltic Federal University, Kaliningrad, Russia

² Lomonosov Moscow State University, Moscow, Russia

³ University of Antwerp, Antwerp, Belgium

* yurgusguss@mail.ru

Key words: mitochondria; mtDNA; evolution; mutational spectrum; birds

Motivation and Aim: Mitochondria is the powerbase of the cells with its own DNA molecule (further mtDNA). During replication mtDNA becomes vulnerable to different mutagens. Replication in mtDNA strands proceeds unevenly [1]. Protein-coding genes, close to the origin of replication (COX1, COX2) spend less time in single-stranded condition than ones farther from it (CytB) on heavy strand [2]. Exposed genes in single stranded states become easy targets for different internal mutagenic factors, including spontaneous deamination or even oxidative damage. The result of this mutagenic process is A>G substitution on the heavy strand. It was shown for different groups of animals that this mutation is connected with different life-history traits. Long-lived mammals demonstrate increased frequencies of A>G transitions compared to the short-lived ones [3]. Also there is a high difference in mutation frequencies of A>G between fishes inhabiting cold and warm water [4]. Armed by this knowledge we decided to analyse mtDNA mutagenesis in birds and potential factors affecting it.

Methods and Algorithms: To study the influence of life-history traits in different birds on their mtDNA mutational spectrum, we collected all available information. First of all we analysed mtDNA RefSeq data for 766 bird species. For each species we calculated a specific metric to study mtDNA nucleotide shift due to A>G mutation – GhAhSkew by using the following formula: $GhAhSkew = (G-A)/(G+A)$. In this formula G and A is the amount of guanine and adenine in four-fold neutral positions. We used it in each analysis to study mtDNA mutagenesis. Moreover we added ecological data from the AVONET database and deployed a specific phenotype classification of birds based on ability to fly and dive with the help of Birds of the World database [5, 6]. To find primary associations between mtDNA mutagenesis and life-history traits we used t-test, PCA, U-test and linear models. To study phylogenetic signals in the observed associations we used specific methods such as PGLS and reconstruction of ancestral states.

Results: By using simple statistical methods we discovered pretty interesting results for RefSeq data. Similar to mammals, birds have a gradient of GhAhSkew in protein-coding genes depending on their time being single stranded during replication [3]. COX1 and COX2 have low GhAhSkew compared to CytB. Having such results we decided to compare each group of animals. We saw that birds have around 1.5 times higher GhAhSkew than mammals. This result drew us to a thought that birds' mutagenesis speed was much higher than mammals. In order to understand the reasons behind such intense mutagenesis, we started to compare our metric GhAhSkew between different birds' life-

history traits. We observed that birds with only one trait had low GhAhSkew – losing ability to fly. Primarily for *Palaeognathae* species. But there were many traits of birds which increased GhAhSkew: ability to dive (especially for *Sphenisciformes* and *Anseriformes*), far migration, thermo-neutral zone. To our surprise, there was no connection between GhAhSkew and many metrics, like body mass for example. After all these discoveries we decided to analyse the phylogenetic value of them. Due to the data limits we were only able to look at ability to fly/dive phenotypes. For all those phenotypes PGLS showed statistically significant results for GhAhSkew and losing ability to fly (p-value < 0.05) and GhAhSkew and getting ability to dive (p-value < 0.05). Both results also had high Pagel's lambda value ($\lambda > 0.9$) which shows strong phylogenetic inertia for the GhAhSkew metric. For those phenotypes we also calculated PGLS with a comparative species-specific A>G spectrum, which is based on our RefSeq data. We found no correlations for A>G frequencies and any phenotypes. We made a hypothesis that birds with our phenotypes of interest are under stabilizing selection in order to save last adenines.

Conclusion: We showed that birds mtDNA A>G mutation rate is quite high, even compared to mammals. It led to increased guanine enrichment in modern birds mtDNA with specific phenotypes: ability to dive, long-distance migration and thermo-neutral zone. But there are also factors for lowering mutagenesis – losing ability to fly. We discovered that all mutagenesis effects emerged at the level of ancestors, and nowadays some bird species are under stabilizing selection. For further research, we have expanded our data by gathering mtDNA information from the MIDORI2 database for more than 7000 bird species [7]. In addition we also got A>G spectra for our birds database: one for our 766 RefSeqs and another for 789 birds by using the NeMu pipeline [8].

Funding: The study is supported by RSF grant (No. 21-75-20143).

References

1. Falkenberg M. Mitochondrial DNA replication in mammalian cells: overview of the pathway. *Essays Biochem.* 2018;62(3):287-296. doi 10.1042/EBC20170100
2. Ju Y.S., Alexandrov L.B., Gerstung M. et al. Origins and functional consequences of somatic mitochondrial DNA mutations in human cancer. *Elife.* 2014;3:e02935. doi 10.7554/eLife.02935
3. Mikhailova A.G., Mikhailova A.A., Ushakova K. et al. A mitochondria-specific mutational signature of aging: increased rate of A > G substitutions on the heavy strand. *Nucleic Acids Res.* 2022;50(18):10264-10277. doi 10.1093/nar/gkac779
4. Mikhailova A. et al. A mitochondrial mutational signature of temperature in ectothermic and endothermic vertebrates. *bioRxiv.* 2021. doi 10.1101/2020.07.25.221184
5. Tobias J.A., Sheard C., Pigot A.L. et al. AVONET: morphological, ecological and geographical data for all birds. *Ecol Lett.* 2022;25:581-597
6. Billerman S.M., Keeney B.K., Rodewald P.G., Schulenberg T.S. (Eds.). *Birds of the World*. Cornell NY, USA: Laboratory of Ornithology, Ithaca, 2022
7. Leray M., Knowlton N., Machida R.J. MIDORI2: A collection of quality controlled, preformatted, and regularly updated reference databases for taxonomic assignment of eukaryotic mitochondrial sequences. *Environmental DNA.* 2022;4(4):894-907. doi 10.1002/edn3.303
8. Efimenko B., Popadin K., Gunbin K. NeMu: a comprehensive pipeline for accurate reconstruction of neutral mutation spectra from evolutionary data. *bioRxiv* 2023. doi 10.1101/2023.12.13.571433