Chris Deardeuff
1236847562
Dr. Nate Upham
BIO-549 Phylogenetic Biology & Analysis
2025-12-09

Background

This Project analyzed and performed a continuous character trait evolution analysis for the family *Spheniscidae*, commonly referred to as penguins. They are part of the order Sphenisciformes, and are flightless birds adapted to live near, and operate in, aquatic environments. *Spheniscidae* often have counter-shading to assist with camouflaging in their aquatic habitats, and their wings are paddle like[1]. The majority of *Spheniscidae* live only in the southern hemisphere[2]. Some of the smaller species of penguins occupy temperate or even tropical climates, while the larger penguins are mostly in the cooler subantarctic regions. There are eighteen extant species composing six genera recognized by the International Ornithologists' Union[3].

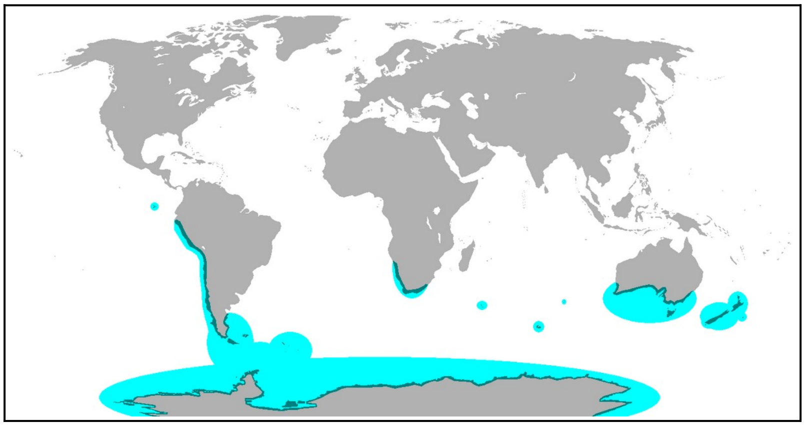

Figure 1: Breeding range of all penguin species

Spheniscidae was chosen for this project because they have a decently varied range in body mass between species and body mass has critical influence on their ecology from dive depth, to thermoregulation, to their ability to forage effectively[4]. With these attributes being indicators of their success in an ecosystem[5], this analysis will look at if the evolution of body mass has been neutral, or if evidence of a stabilizing selection is present, providing further evidence of its importance.

Methods and Materials

The base tree being used for this project is a Maximum Clade Credibility (MCC) tree from Vertlife[6]. Overall there are 18 unique penguin species, with an outgroup species Giant Southern Petral. Using this tree, and traits for *Spheniscidae* that were obtained through EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals[7], a continuous character trait evolution analysis was performed. The table below shows the Body Mass value for each of the species in the analysis.

Table 1: Table containing all body mass trait data for each penguin species in the analysis.

Species ID	Name	Common Name	Body Mass Average (grams)	Body Mass Average ln(g)
5413	*Aptenodytes patagonicus*	King Penguin	11731.07	9.37
5414	*Aptenodytes forsteri*	Emperor Penguin	33569.33	10.42
5415	*Pygoscelis papua*	Gentoo Penguin	5932.95	8.69
5416	*Pygoscelis adeliae*	Adelie Penguin	4847.67	8.49
5417	*Pygoscelis antarcticus*	Chinstrap Penguin	4146.08	8.33
5418	*Eudyptes chrysocome*	Southern Rockhopper Penguin	2327.85	7.75
5422	*Eudyptes pachyrhynchus*	Fiordland Penguin	3904.88	8.27
5423	*Eudyptes robustus*	Snares Penguin	3038.02	8.02
5424	*Eudyptes sclateri*	Erect-crested Penguin	5888.95	8.68
5425	*Eudyptes chrysolophus*	Macaroni Penguin	4485.08	8.41
5426	*Eudyptes schlegeli*	Royal Penguin	4242.64	8.35
5427	*Megadyptes antipodes*	Yellow-eyed Penguin	5326.24	8.58
5428	*Eudyptula minor*	Little Penguin	1107.82	7.01
5433	*Spheniscus demersus*	African Penguin	3130.11	8.05
5434	*Spheniscus humboldti*	Humboldt Penguin	4366.39	8.38
5435	*Spheniscus magellanicus*	Magellanic Penguin	4105.1	8.32
5436	*Spheniscus mendiculus*	Galapagos Penguin	1921.86	7.56

The trait data referenced above was formatted into a nexus file with the body mass average being log transformed to have a more uniform distribution. Both the *Macronectes giganteus* (Giant Southern Petrel) and *Eudyptes moseleyi* (Northern Rockhopper Penguin) tips were dropped from the tree before analysis. The Giant Southern Petrel was the outgroup taxon while the Northern Rockhopper Penguin was not available in the dataset from Elton Traits.

A Continuous-Character Evolution analysis was performed using RevBayes. This analysis will use two different models, the Brownian Motion (BM) model and Ornstein-Uhlenbeck (OU) model. These two models will be compared with a Reverse Jump – Monte Carlo Markov Chain (RJ-MCMC) and calculating the Bayes Factor (BF). The BM model will assume that the body mass trait evolves randomly and with a drift parameter σ^2. The OU model will include extra parameters which will pull the trait toward an optimum. This will represent body mass being a stabilizing selection or constraint on the evolution of penguins. The OU model's parameters include: α – representing the strength of the selection. Θ – The optimum value, or target. And lastly σ^2, the same parameter as in the BM model, representing stochastic random drift. The tutorials being referenced for this analysis are the Simple Ornstein-Uhlenbeck Model[8] tutorial and Simple Brownian Model[9] available through RevBayes. All scripts, data, and output can be found in a GitHub repository here.

Results

The Brownian Model ran for 50,000 iterations and had an effective sample size (ESS) of 8597.7 for σ^2 indicating the chain mixed well. The Posterior Mean Rate for σ^2 was 0.0429 and the 95% HPD Interval was [0.0174, 0.0749]. These show that with this model estimated instantaneous rate of squared change in ln(Body Mass) for penguins is approximately 0.043. The following table is from the log file of the BM analysis.

Table 2: Table with log values from the Brownian Model Analysis

Summary Statistic	Posterior	Likelihood	Prior	Sigma2
Mean	-13.55	-15.37	1.82	0.04
Stderr of mean	0.01	0.01	0	0
Stdev	0.74	0.62	0.35	0.02
Variance	0.55	0.39	0.12	0
Median	-13.24	-15.12	1.84	0.04
Value Range	[-17.5412, -12.9576]	[-20.4006, -14.8948]	[0.9169, 2.8593]	[0.0143, 0.0999]
Geometric Mean	n/a	n/a	1.79	0.04
95% HPD interval	[-15.2101, -12.9576]	[-16.7123, -14.8948]	[1.101, 2.4645]	[0.0174, 0.0749]
Auto-correlation time (ACT)	1.04	1.03	1.04	1.05
Effective sample size (ESS)	8625.4	8719.6	8643.6	8597.7
Number of samples	9002	9002	9002	9002

Table 3: Table with log values from the Ornstein-Uhlenbeck Model Analysis

Summary Statistic	Posterior	Likelihood	Prior	Alpha	P_Th	Sigma2	T_Half	Theta
Mean	-19.16	-16.71	-2.45	0.09	0.79	0.33	8.5	9.3
Stderr of mean	0.02	0.01	0.01	0	0	0	0.02	0
Stdev	1.55	1.15	0.72	0.03	0.05	0.07	2.02	0.45
Variance	2.41	1.33	0.51	0	0	0.01	4.09	0.21
Median	-18.81	-16.42	-2.28	0.08	0.79	0.31	8.65	9.36
Value Range	[-29.1135, -16.9916]	[-24.5398, -15.2292]	[-8.6486, -1.2586]	[0.0417, 0.3513]	[0.6274, 0.9516]	[0.1799, 0.8897]	[1.9732, 16.6279]	[7.8049, 9.9998]
Geometric Mean	n/a	n/a	n/a	0.08	0.79	0.32	8.23	9.29
95% HPD Interval	[-22.2239, -17.0246]	[-18.9741, -15.2418]	[-3.9065, -1.4812]	[0.0511, 0.1425]	[0.7134, 0.8941]	[0.203, 0.4719]	[4.2357, 12.0101]	[8.5187, 9.9998]
Auto-Correlation Time (ACT)	1.01	1	1.05	1.04	1.04	1.04	1.17	1.04
Effective Sample Size (ESS)	8896.8	9002	8546.3	8621.6	8694.6	8644.8	7674.1	8653.2
Number of Samples	9002	9002	9002	9002	9002	9002	9002	9002

The Ornstein-Uhlenbeck analysis also ran for 50,000 iterations and had large ESS values indicating it mixed well. The Ɵ parameter's posterior mean was 9.30. This represents the optimal body mass for penguins and untransformed equals 10,986 grams. The posterior mean of α was 0.088 and had a 95% HPD interval of [0.052, 0.144], this is informative as it shows a pull towards the optimum that was never at zero.

Using an RJ-MCMC analysis, the two models can be compared using Bayes Factor (BF)[10]. The calculation for BF is:

$$\ln(\mathrm{BF}) = \ln\left(\frac{P_{\mathrm{post}}(\mathrm{OU})}{P_{\mathrm{post}}(\mathrm{BM})}\right) - \ln\left(\frac{P_{\mathrm{prior}}(\mathrm{OU})}{P_{\mathrm{prior}}(\mathrm{BM})}\right)$$

With the values from the analysis plugged in:

$$\ln(\mathrm{BF}) = \ln\left(\frac{0.9002}{0.0998}\right) - \ln\left(\frac{0.5}{0.5}\right)$$

$$\ln(\mathrm{BF}) = \ln(9.0200) - \ln(1.0)$$

$$\ln(\mathrm{BF}) \approx 2.1994$$

This value provides strong support for the OU model and will be discussed more later.

Discussion

With a Bayes Factor value of ln(2.1994), squarely in support of the OU model[11], the analysis supports that body mass is a stabilizing evolutionary factor for penguins. This finding goes along with the recent research into body mass for penguins and its importance and benefits. The optimum value of ln(9.30) fits just above the the median of 8.71 of current penguin body mass as seen in figure 2. Further evidence of larger body mass being beneficial for penguins foraging, diving depth, and thermal regulation[12].

For future analysis adding more character traits such as feeding characteristics, and geographical ranges, could further provide evidence on the impact of body mass, and how penguins interact with their ecosystems and what are the main

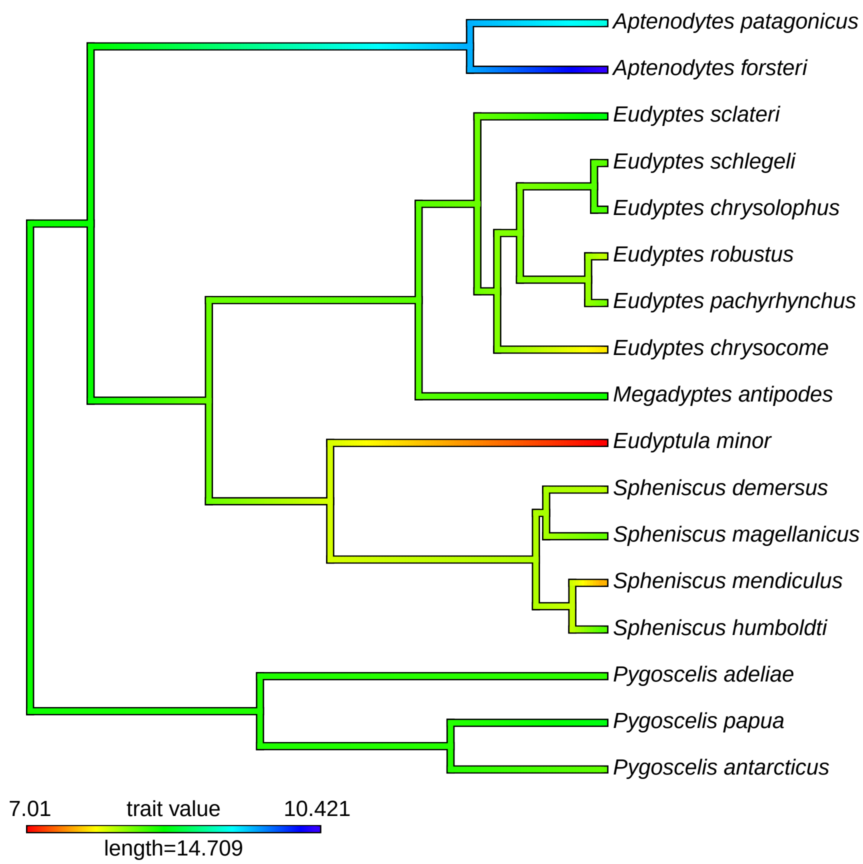

Figure 2: Character mapping of the body mass in ln(g) to the phylogeny for Spheniscidae.

drivers of their evolution. Also adding in extinct taxa with estimated body mass could be interesting to see how much it changes the optimum for the OU analysis.

References

Figures

1. File:Penguin range.png - Wikimedia Commons. (2022). Wikimedia.org. [Wikimedia]
2. File: Character mapping of the body mass in ln(g) to the phylogeny for Spheniscidae.

Tables

1. Data extracted from Jetz, W. (2016). EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals. *Figshare*. [Figshare]
2. File: BM_datatable.txt - Table with log values from the Brownian Model Analysis, [LINK HERE]
3. File: OU_datatable.txt - Table with log values from the Ornstein-Uhlenbeck Model Analysis, [LINK HERE]

References

1. Prevost, J. (2020). Penguin - Natural history | Britannica. In *Encyclopædia Britannica*. https://www.britannica.com/animal/penguin/Natural-history
2. Prevost, J. (2020). Penguin - Natural history | Britannica. In *Encyclopædia Britannica*. https://www.britannica.com/animal/penguin/Natural-history
3. Gill F, D Donsker & P Rasmussen (Eds). 2025. IOC World Bird List (v15.1). https://www.worldbirdnames.org/bow/loons/
4. Sato, K., Naito, Y., Kato, A., Niizuma, Y., Watanuki, Y., Charrassin, J. B., Bost, C.-A. ., Handrich, Y., & Le Maho, Y. (2002). Buoyancy and maximal diving depth in penguins: do they control inhaling air volume? *Journal of Experimental Biology*, *205*(9), 1189–1197. https://doi.org/10.1242/jeb.205.9.1189
5. Masud, M. H., & Dabnichki, P. (2024). Biomechanical analysis of little penguins' underwater locomotion from the free-ranging dive data. *Biology Open*, *13*(5). https://doi.org/10.1242/bio.060244
6. Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., & Mooers, A. O. (2012). The global diversity of birds in space and time. *Nature*, *491*(7424), 444–448. https://doi.org/10.1038/nature11631
7. Wilman, Hamish; Belmaker, Jonathan; Simpson, Jennifer; de la Rosa, Carolina; Rivadeneira, Marcelo M.; Jetz, Walter (2016). EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals. Wiley. Collection. https://doi.org/10.6084/m9.figshare.c.3306933.v1
8. *RevBayes: Simple Ornstein-Uhlenbeck Models*. (2022). Github.io. https://revbayes.github.io/tutorials/cont_traits/simple_ou.html
9. *RevBayes: Simple Brownian Rate Estimation*. (2022). Github.io. https://revbayes.github.io/tutorials/cont_traits/simple_bm

10. *RevBayes: Simple Ornstein-Uhlenbeck Models*. (2022). Github.io. https://revbayes.github.io/tutorials/cont_traits/simple_ou.html
11. Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, *90*(430), 773. https://doi.org/10.2307/2291091
12. Sato, K., Naito, Y., Kato, A., Niizuma, Y., Watanuki, Y., Charrassin, J. B., Bost, C.-A. ., Handrich, Y., & Le Maho, Y. (2002). Buoyancy and maximal diving depth in penguins: do they control inhaling air volume? *Journal of Experimental Biology*, *205*(9), 1189–1197. https://doi.org/10.1242/jeb.205.9.1189