

Modeling multivariate trait evolution using mixed Gaussian phylogenetic models

Venelin Mitov^{1,2}, Krzysztof Bartoszek³, Georgios Asimomitis¹, Tanja Stadler^{1,2}

1 Department of Biosystems, Science and Engineering, Swiss Federal Institute of Technology Zurich, Basel, Switzerland

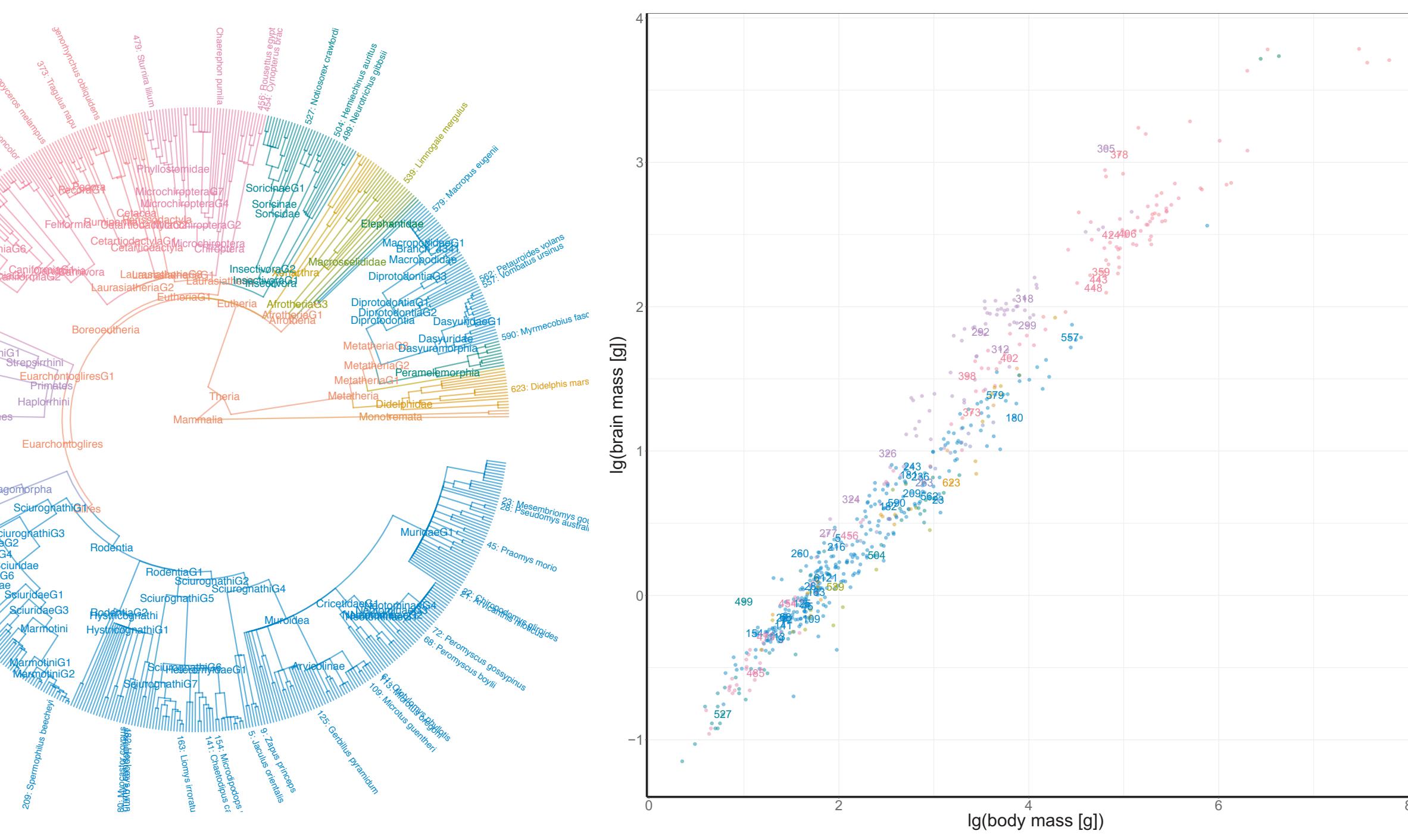
2 Swiss Institute of Bioinformatics

3 Linköping University, Sweden

Gaussian phylogenetic models like Brownian motion (BM) and Ornstein-Uhlenbeck (OU) are the work-horse of modeling continuous trait evolution. However, these models fit poorly to big phylogenetic trees, because they cannot account for non-homogenous evolutionary processes. This issue has been addressed by introducing shifts at predefined or inferred points on the tree. In most current implementations, though, these shifts are "intra-model", meaning that they allow a shift in one or two parameters of the model, such as the selective optimum of an OU process, keeping all other parameters "global" for the entire tree. The more ambitious task of finding "inter-model" shifts involving all parameters in a model as well as different types of models has not been addressed due to its computational complexity. Here, we propose a reasonable compromise between the generality of the inter-model shift problem and the computational feasibility of searching through the space of shifts and model mappings. We illustrate our approach with an analysis of previously published comparative dataset from Mammals.

Data [1,2]

1. A phylogenetic tree comprising 629 mammal species
2. $\lg(\text{body mass})$ and $\lg(\text{brain mass})$ data for each species



Question: How did the two traits evolve through time?

Methods

1. Mixed Gaussian Phylogenetic Models

Assume that the traits have co-evolved along the branches of the tree following a continuous time Gaussian Markov process, satisfying the following properties:

- (1.1) every two branches evolve independently after their most recent common ancestor;
- (1.2) the distribution of the trait-vector at time t conditional on an ancestor value at time s is Gaussian, such that
 - (a) the expectation-vector depends linearly on the ancestral value;
 - (b) the variance-covariance matrix is independent of the ancestral trait.
- (1.3) At any point on a branch a model shift can take place, the new model inheriting the current trait-value of the previous model.

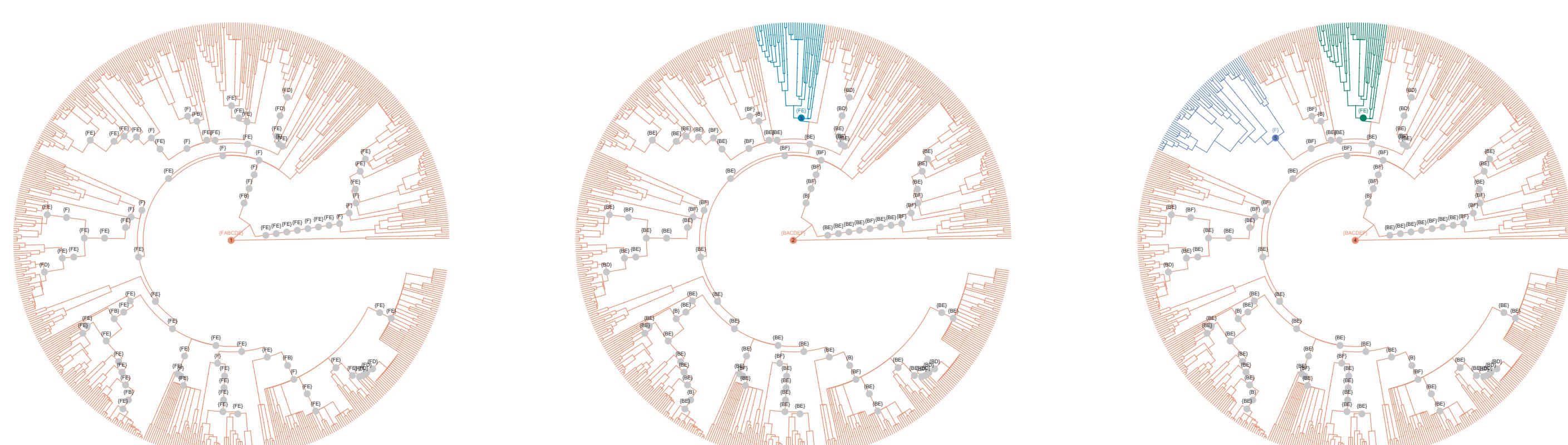
For the mammal dataset, we used the following six models:

- BM_A : A Brownian motion (BM) with diagonal variance covariance matrix (Σ) – uncorrelated traits;
- BM_B : A BM with non-diagonal Σ – correlated traits;
- OU_C : An Ornstein-Uhlenbeck (OU) with diagonal selection strength matrix (H) and diagonal Σ ;
- OU_D : An OU with diagonal H and non-diagonal symmetric Σ ;
- OU_E : An OU with non-diagonal symmetric H and non-diagonal symmetric Σ ;
- OU_F : An OU with non-diagonal symmetric H and non-diagonal symmetric Σ ;

2. Linear time likelihood calculation using a quadratic polynomial formulation [3,4]

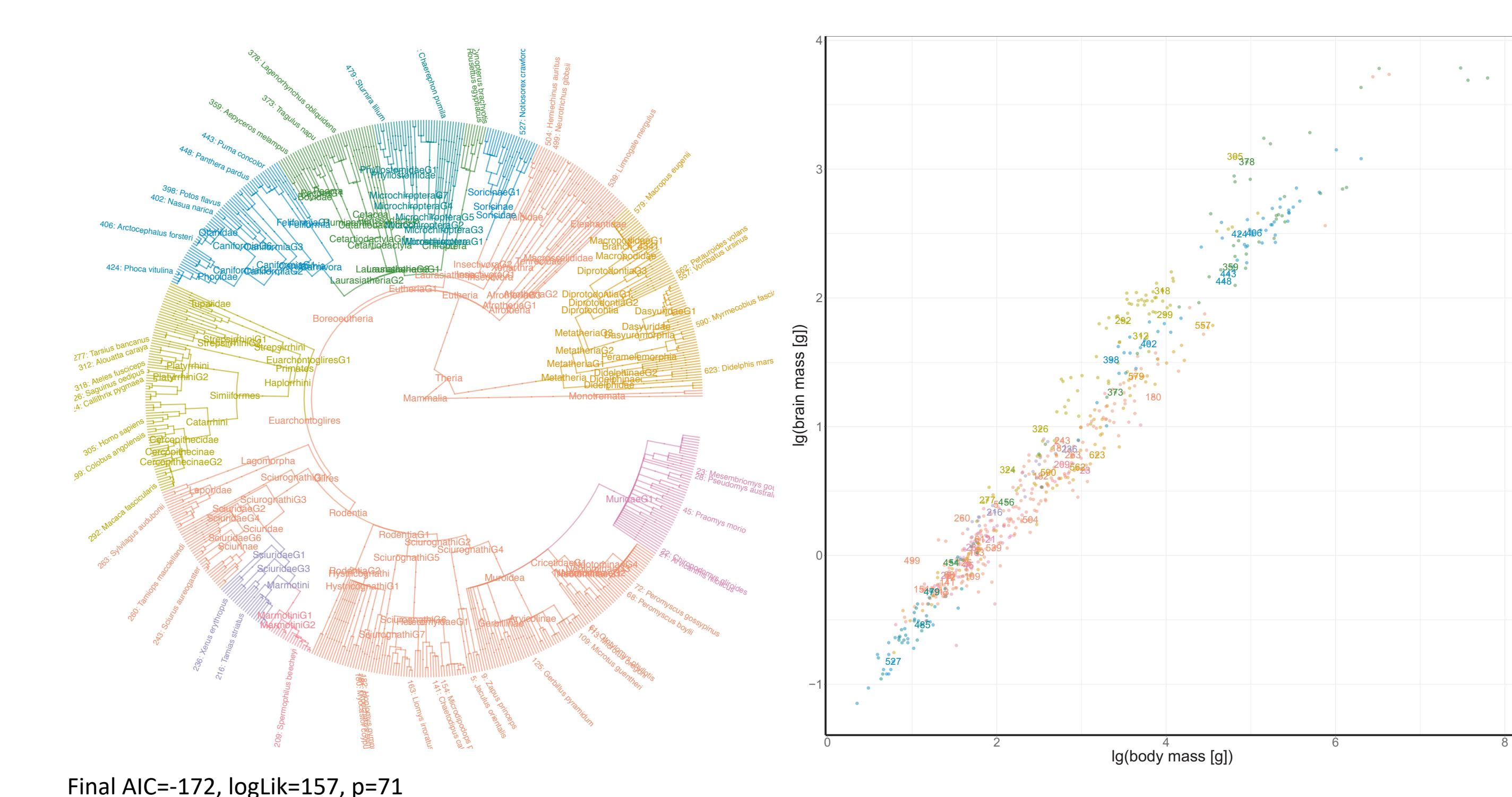
3. Parallel stepwise AIC procedure for finding optimal shifts and models mappings:

(1) AIC=485, IL=-230, p=12 (2) AIC=-28, IL=31, p=17 (4) AIC=-63, IL=59, p=28

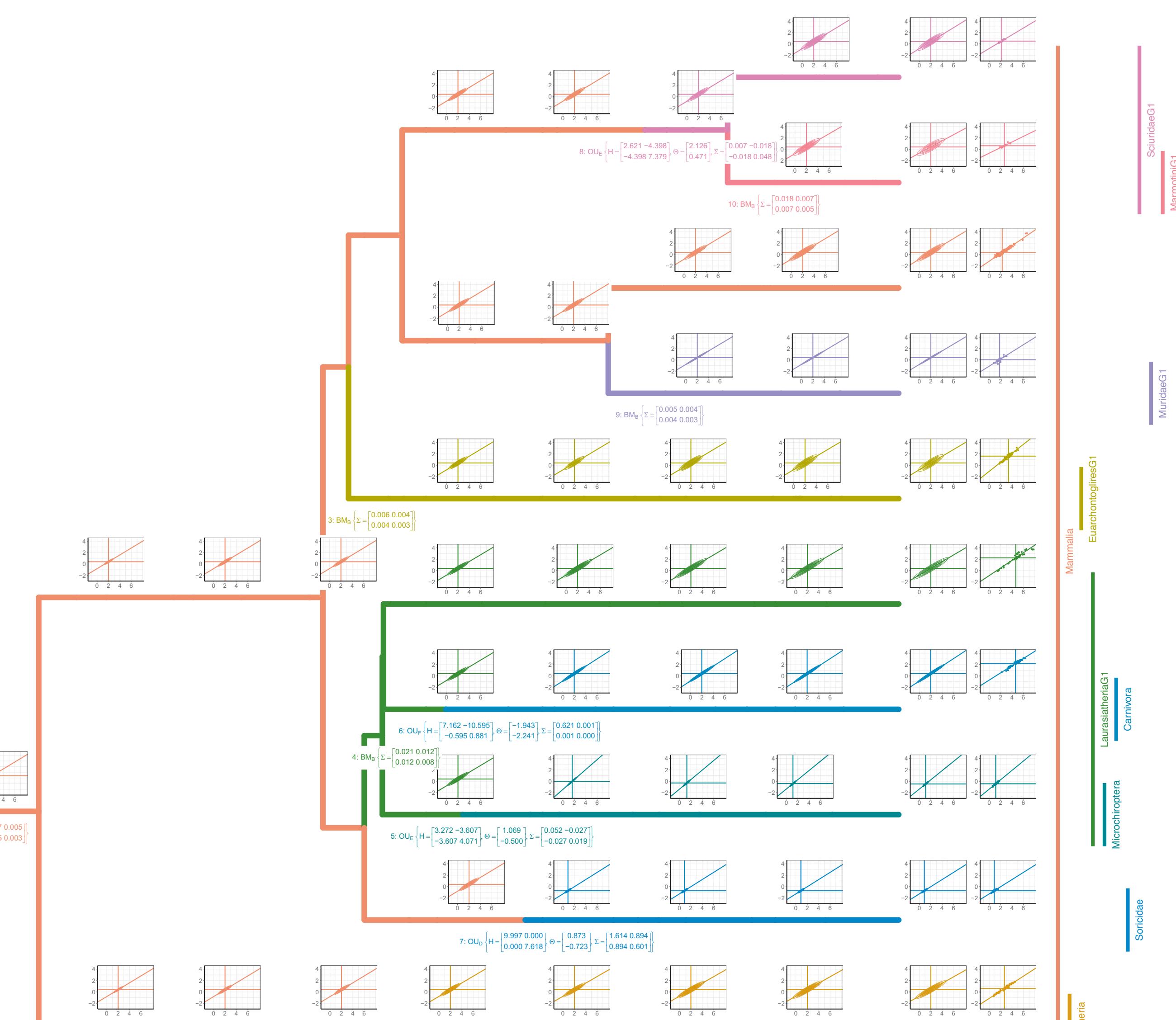


Results

1. Inferred shifts show an approximate match with mammal orders



2. Inferred models and "predicted" trait distributions through time:



References

- [1] Bininda-Emonds, O. R. P., Cardillo, M., Jones, K. E., MacPhee, R. D. E., Beck, R. M. D., Grenyer, R., et al. (2007). The delayed rise of present-day mammals. *Nature*, 446(7135), 507–512. <http://doi.org/10.1038/nature05634>
- [2] Boddy, A. M., McGowen, M. R., Sherwood, C. C., Grossman, L. I., Goodman, M., & Wildman, D. E. (2012). Comparative analysis of encephalization in mammals reveals relaxed constraints on anthropoid primate and cetacean brain scaling. *Journal of Evolutionary Biology*, 25(5), 981–994. <http://doi.org/10.1111/j.1420-9101.2012.02491.x>
- [3] Mitov, V., & Stadler, T. (2017). POUMM: An R-package for Bayesian Inference of Phylogenetic Heritability. preprint <http://doi.org/10.1101/115089>
- [4] Mitov, V., Bartoszek, K., Stadler, T. (2018). Manuscript in preparation.