# Life history traits affect the magnitude of male mutation bias across 32 mammalian genomes

**Melissa A. Wilson Sayres[1,2,3], Chris Venditti[4], Francesca Chiaromonte[2,3,5], Mark Pagel[4], and Kateryna D. Makova[1,2,3]**

[1] Department of Biology, The Pennsylvania State University, University Park, Pennsylvania, USA
[2] Center for Comparative Genomics and Bioinformatics, The Pennsylvania State University, University Park, Pennsylvania, USA
[3] Integrative Biosciences Program, The Pennsylvania State University, University Park, Pennsylvania, USA
[4] School of Biological Sciences, University of Reading, Reading, Berkshire, RG6 6BX, UK
[5] Department of Statistics, The Pennsylvania State University, University Park, Pennsylvania, USA

## Abstract

Male mutation bias theory predicts that the mutation rate in males is often higher than in females because male gametes, sperm, undergo significantly more rounds of replication than female gametes, eggs (Miyata, *et. al.* 1987, **Figure 1**). Male mutation bias has been observed in mammals, birds, fish, and even plants. Curiously, however, *estimates of the magnitude of male mutation bias vary substantially across species, and even within the same species.* There are two explanations for this. First, differences between estimates from the same species can be explained by regional variation in genome architecture. Not all nucleotide substitutions are affected equally by errors in replication (*e.g.* CpG vs. nonCpG sites). Furthermore, many genomic factors (*e.g.* repetitive elements, GC content, recombination rate) influence mutation rates regionally across the genome and, when not accounted for, can skew estimates of male mutation bias. Second, variations observed across species may be influenced by differences in life history traits, specifically metabolic rate, sexual selection, and generation time. Male mutation bias is expected to be influenced by metabolic rate because sperm not only live in a more reactive oxygen species rich environment, they are also more susceptible to mutations through oxidative stress than eggs. Male mutation bias might become elevated with stronger post-copulatory sexual selection; males in species where sperm from multiple males compete to fertilize eggs produce more sperm, potentially at the expense of a higher mutation rate, than males in species without competition. Additionally, species with shorter generation times might experience less male mutation bias because their sperm undergo fewer rounds of replication before conception.

Few studies have investigated the factors that influence variation in the magnitude of male mutation bias across multiple species using more than a subset of any genome. Utilizing the 32 eutherian mammal genome sequences we are able to investigate male mutation bias on a genome-wide scale across mammalian taxa with diverse life history traits. **We ask which life history traits affect the magnitude of male mutation bias observed in mammals.** To answer this question we collected literature on life history traits for all 32 mammals, filtered whole-genome alignments of factors known to influence substitution rates regionally, and computed global and context-dependent substitution rates. Then, after accounting for phylogenetic dependence, we developed a model to define how variations in life history traits affect variations in the magnitude of male mutation bias. We found that representatives of three major life history traits (metabolic rate, generation time, and sexual selection) all affect the magnitude of male mutation bias, explaining at least 70% of the variation in the magnitude of male mutation bias observed across these diverse mammals. **Our results corroborate and expand upon years of previous research, and support the significant effect of life history traits on genome evolution.**

## Materials and methods

$$A = 1/2(\mu_M) + 1/2(\mu_F)$$
$$X = 1/3(\mu_M) + 2/3(\mu_F)$$
$$Y = \mu_M$$

**Figure 1. Relative time spent in the male and female germlines.** Autosomes (A) spend half of their time in males and half in females, while the sex chromosomes are split disproportionately; X spends 2/3 of its time in females while Y is found exclusively in males. Thus, replication-dependent mutation rates ($\mu_M$ and $\mu_F$), should affect substitution rates observed on X, Y and A in the proportions listed in the figure.

| Life history trait | Can be approximated by: |
|---|---|
| **Generation Time** | - Age at Sexual Maturity (SexMat)<br>- Lifespan<br>- Inter-litter Interval (Inter) |
| **Metabolic Rate** | - Basal Metabolic Rate (BMR)<br>- Body Temperature (Temp)<br>- Mass |
| **Sperm Competition** | - Testis to Body Mass Ratio (TBR)<br>- Mating Pattern (Mating) |

**Table 1. Biological features used to approximate life history traits.** Literature indicates that the biological features listed in the table to the right may approximate each of the three major life history traits of interest. Abbreviations for each biological feature are in parentheses.

## Results

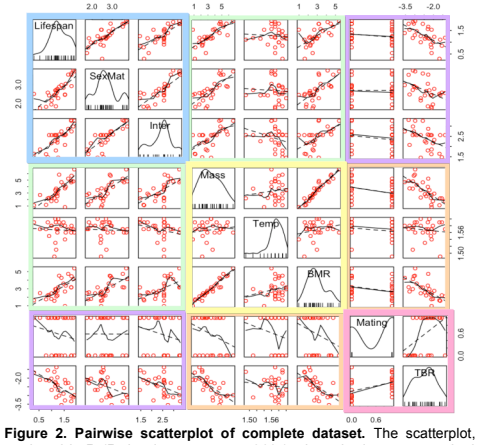


**Figure 2. Pairwise scatterplot of complete dataset.** The scatterplot, produced in R (R development team, *2005*), show the lowess smoothed curves and the regression lines of each pairwise plot of predictors.

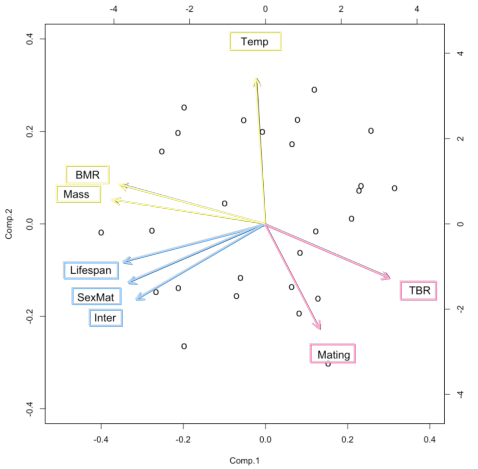


**Figure 3. Principal Component Analysis (PCA) of the 8 predictors.** PCA identifies components that, decreasingly, describe the most variability in the dataset. Biological features for each life history trait group together. Likely, temperature is an outlier because it is difficult to collect consistent estimates of body temperature.

| R^2 | Generation Time | | | Metabolic Rate | | | Sperm Competition | |
|---|---|---|---|---|---|---|---|---|
| | Lifespan | SexMat | Inter | Mass | Temp | BMR | Mating | TBR |
| **0.7062** RCVE | 0.63 | | | | | 0.63 | | 0.45 |
| *P* | *0.0464* | | | | | *<.0001* | | *<.0001* |
| *(VIF)* | *(2.14)* | | | | | *(3.27)* | | *(2.45)* |

**Table 2. Optimal regression model.** The model above explains over 70% of the variation in the data, includes at least one significant predictor representing each life history trait, and all predictors have a VIF < 4. Variance Inflation Factor (VIF) describes the level of co-variation among traits. Higher VIF values indicate more co-variation among traits. In both Tables 2 and 3: red indicates a positive coefficient, blue indicates a negative coefficient.

| R^2 | Generation Time | | | Metabolic Rate | | | Sperm Competition | |
|---|---|---|---|---|---|---|---|---|
| | Lifespan | SexMat | Inter | Mass | Temp | BMR | Mating | TBR |
| **0.7125** RCVE | | | | 0.67 | | 0.84 | 0.088 | |
| *P* | | | | *<.0001* | | *<.0001* | *0.0168* | |
| *(VIF)* | | | | *(3.61)* | | *(3.77)* | *(1.80)* | |
| **0.7118** RCVE | | | | 0.72 | | 0.99 | 0.1050 | |
| *P* | | | | *<.0001* | | *<.0001* | *0.0094* | |
| *(VIF)* | | | | *(3.42)* | | *(3.24)* | *(1.15)* | |
| **0.5363** RCVE | | 0.52 | | 0.13 | | | 0.14 | 0.15 |
| *P* | | *0.0464* | | *0.0039* | | | *0.0039* | *0.0034* |
| *(VIF)* | | *(1.33)* | | *(1.47)* | | | *(1.96)* | *(2.48)* |
| **0.5121** RCVE | | 0.24 | | 0.09 | | 0.39 | 0.11 | |
| *P* | | *0.0110* | | *0.0136* | | *0.0139* | *0.0126* | |
| *(VIF)* | | *(2.45)* | | *(1.46)* | | *(2.76)* | *(1.37)* | |

**Table 3. Top regression models.** Four other models explain more than 50% of the variation in male mutation bias and have predictors with VIF < 4. Generally models including predictors of Mass and BMR are excluded due to high VIF, and although two models here include both, the high correlation between these traits should be considered. Correction for phylogenetic dependence and the regression analyses were conducted in BayesTraits (Pagel, *1999*).

## Mammalian Genomes

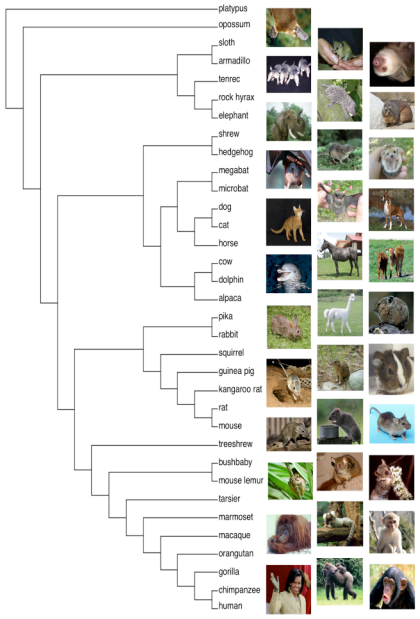


**Figure 4. Mammals in the 44-way species alignment**. The phylogeny of all mammals included in the 44-way species alignment are shown, along with representative photos of each species. Photos follow the branch labels, if read from left to right, starting at the top of the figure.

## Conclusions

- Variation in life history traits explain over 70% of the variation in estimates of male mutation bias (MMB) across mammals.
- Many predictors are highly correlated with one another and likely have compound effects on MMB.
- Generation Time is positively correlated with MMB, supporting hypotheses that increasing time spent in the male germline increases the magnitude of MMB.
- Sperm Competition is generally positively correlated with MMB, supporting hypotheses that as sperm competition intensifies, MMB increases.
- Metabolic Rate is negatively correlated with MMB, suggesting that as Metabolic Rate increases, its mutagenic effect may affect the female germline more severely than the male germline, resulting in a decrease in MMB.
- These results suggest further investigations including detailed studies of male and female germline biochemistry.

## Literature cited

Miyata, *et al.*, Cold Spring Harbor Symposium on Quantitative Biology (*1987*).
Pagel, Nature (*1999*).
R development team (*2005*).

## Acknowledgments

We thank Makova lab members for discussions about this research and gratefully acknowledge the funding for this project provided by an NSF Graduate Research Fellowship to MAWS and by NIH grants RO1 GM072264-05S1 and RO1 GM072264 to KDM. This work was supported in part through instrumentation funded by the National Science Foundation through grant OCI-0821527

## For further information

Please contact *kdm16@psu.edu* or *maw397@psu.edu*. More information on this and related projects can be obtained at: *www.bx.psu.edu/makova_lab*.