Comparing the distribution of fitness effects in great apes

David Castellano, Marie-Curie fellow at Weghorn Lab (CRG)

SRGE Seminar
13th of Sept.

Comparison of the Full Distribution of Fitness Effects of New Amino Acid Mutations Across Great Apes

David Castellano, Moisès Coll Macià, Paula Tataru, Thomas Bataillon and Kasper Munch

GENETICS *Early online September 5, 2019; https://doi.org/10.1534/genetics.119.302494*

Abstract

The distribution of fitness effects (DFE) is central to many questions in evolutionary biology. However, little is known about the differences in DFEs between closely related species. We use more than 9,000 coding genes orthologous one-to-one across great apes, gibbons, and macaques to assess the stability of the DFE across great apes. We use the unfolded site frequency spectrum of polymorphic mutations (n = 8 haploid chromosomes per population) to estimate the DFE. We find that the shape of the deleterious DFE is strikingly similar across great apes. We confirm that effective population size (N_e) is a strong predictor of the strength of negative selection, consistent

In this talk

Introduction

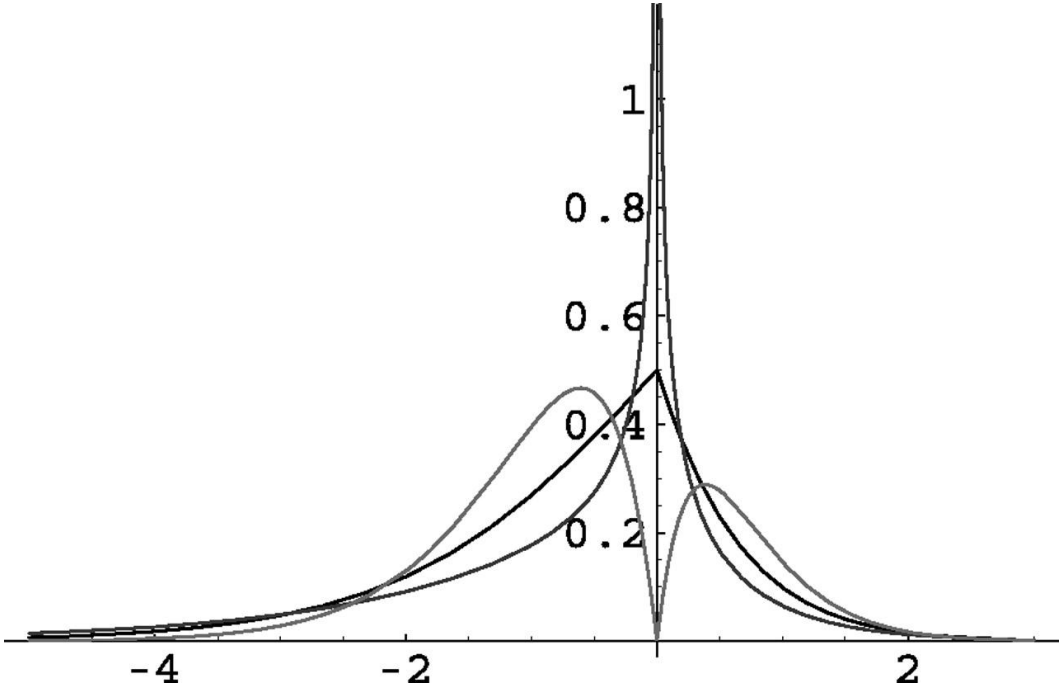
1. What is the distribution of fitness effects (DFE)?
2. Methods to estimate the DFE
    a. Statistical methods to estimate & compare the DFE
    b. polyDFE
3. Previous DFE comparisons and the importance of Fisher's geometrical model (FGM)
4. Great Apes as a genetic system

Aim: Does N_e affect the full DFE all else being equal?

Results & Discussion

Conclusions: Our study demonstrates the simple, but perhaps underappreciated fact that mutational effects are very dynamic even between closely related species.

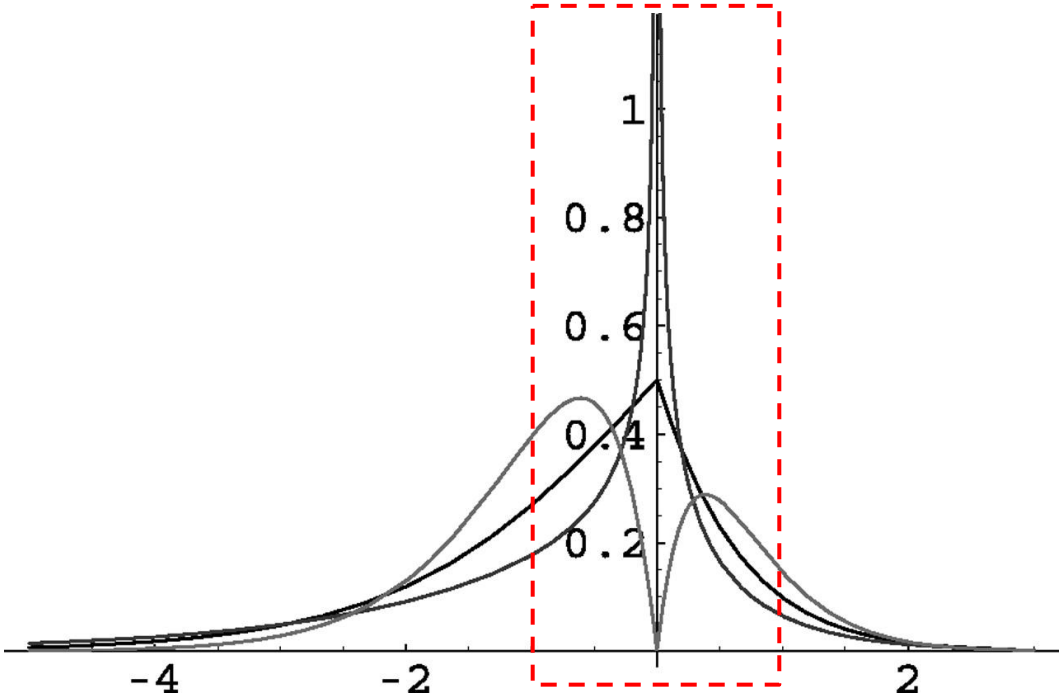
What is the distribution of fitness effects (DFE)?



$S = 2N_e s$

Piganeau & Eyre-Walker 2003

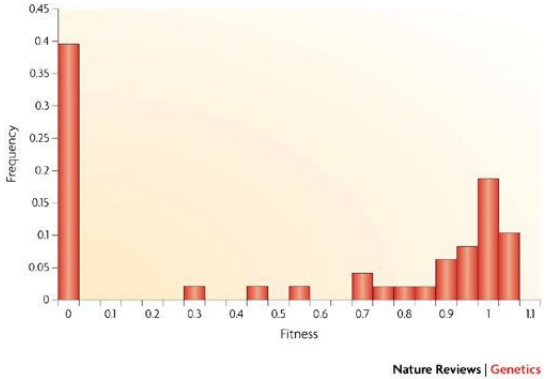
Nearly neutral theory (NNT, Ohta 1973) and the DFE

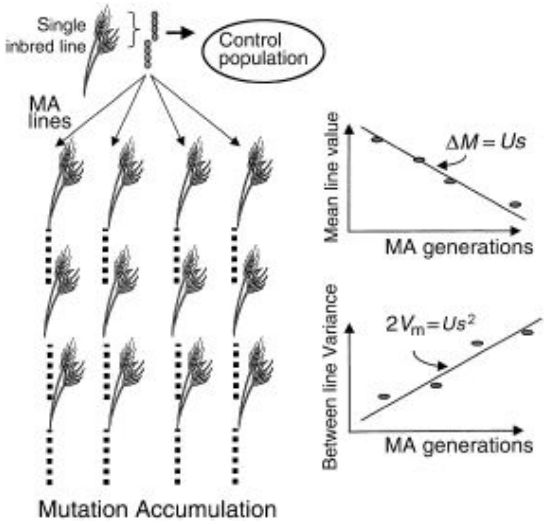


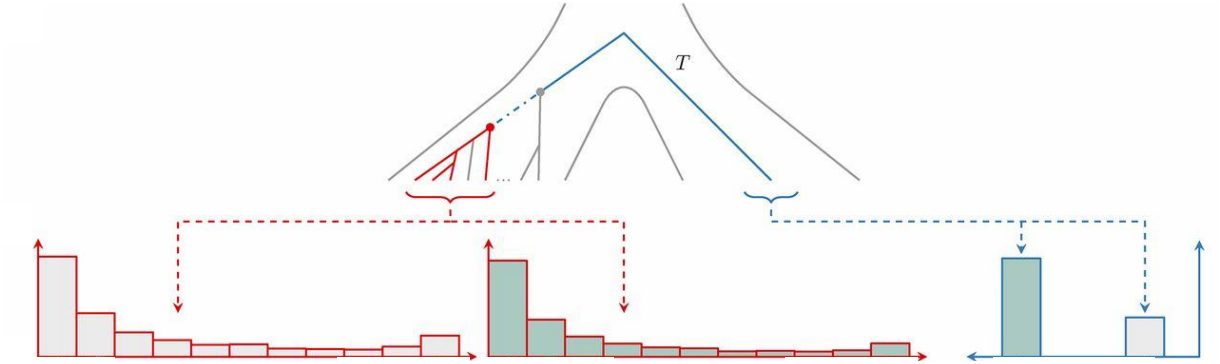
$$S = 2N_e s$$

Methods to estimate the DFE

1. Experimental techniques
    a. Random/Induced mutations and fitness assay
    b. Mutation accumulation experiments
2. Statistical techniques: DNA based





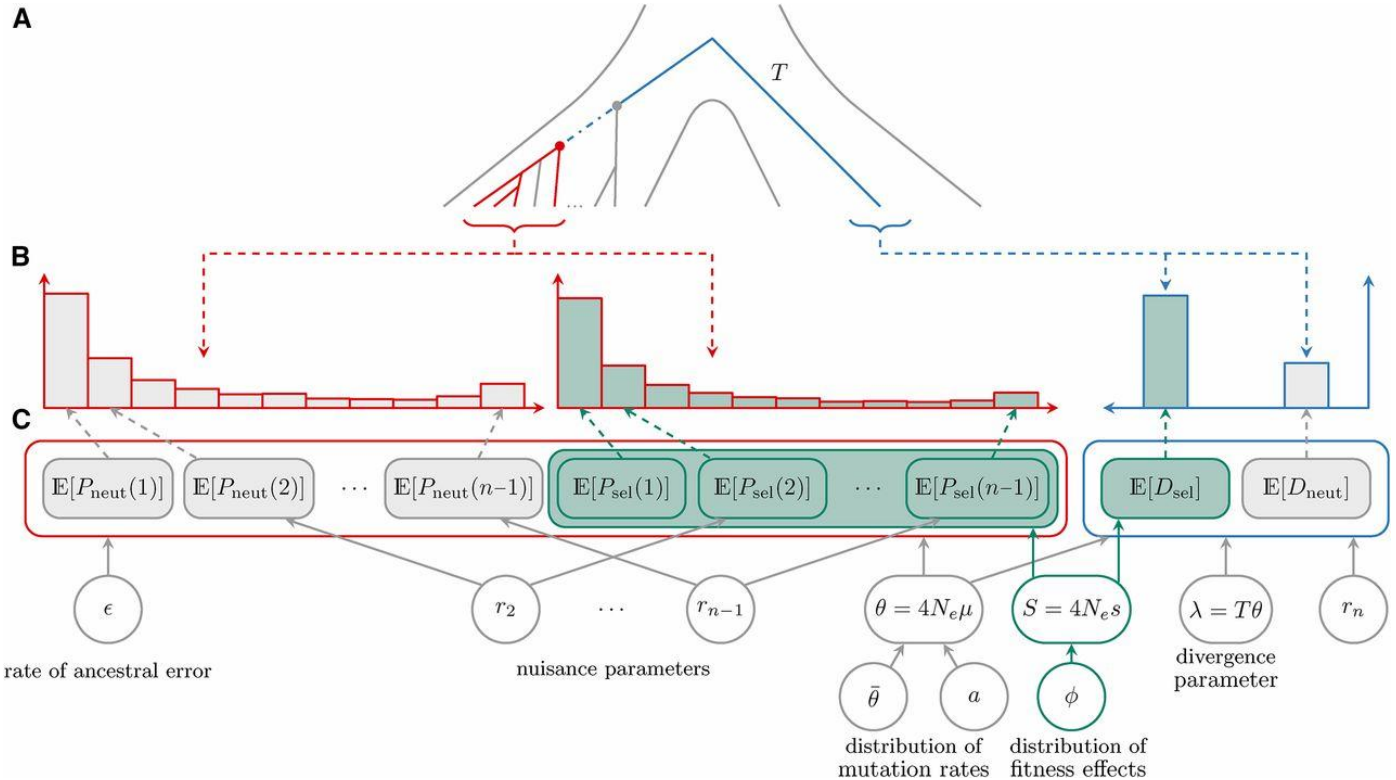


polyDFE, Tataru *et al* **2016**. Genetics; Eyre-Walker & Keightley **2007**. Nature Review Genetics; Bataillon **2000**. Heredity.

Statistical methods to estimate & compare the DFE

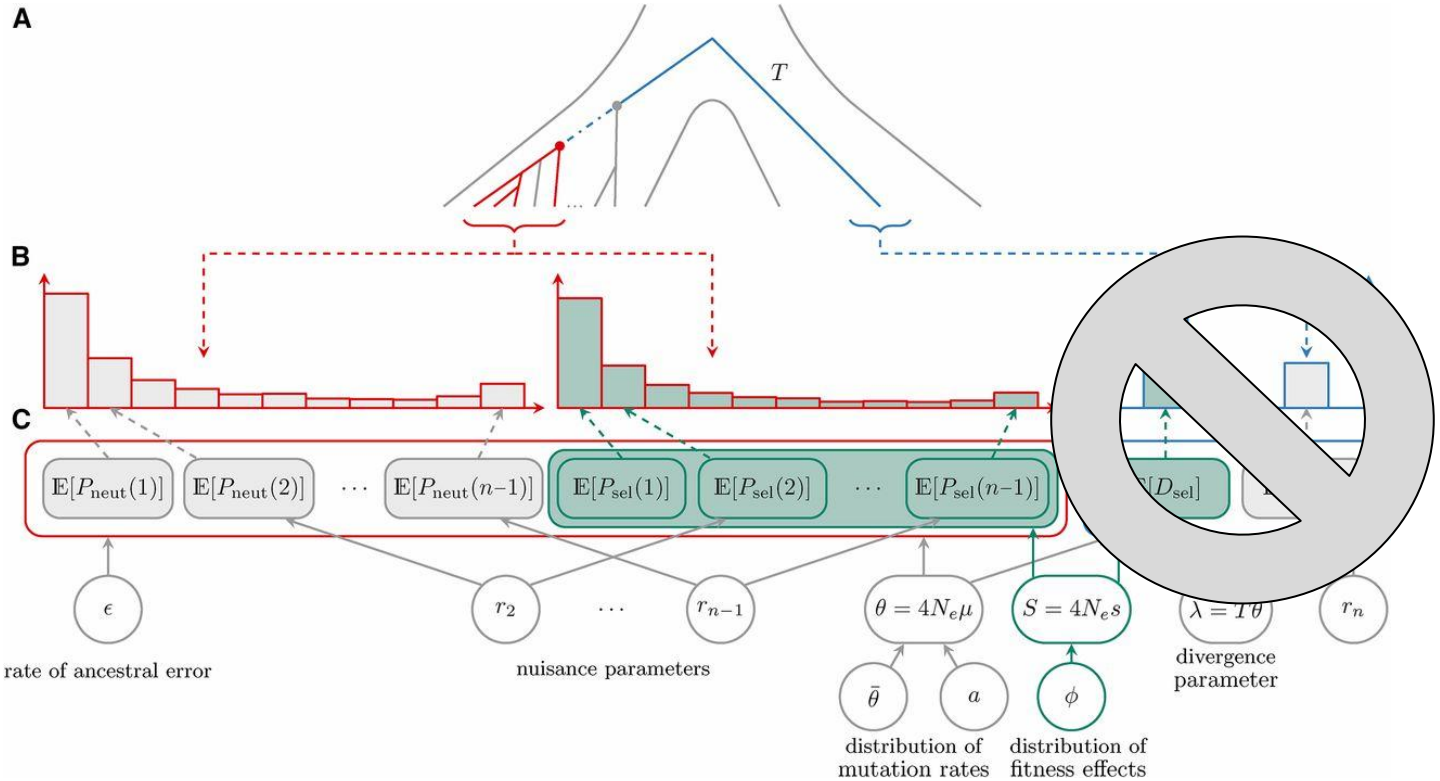
Method (ref)	Deleterious DFE	Beneficial DFE	Demography	Compare
DoFe: Eyre-Walker et al. **2006**. Genetics.	yes	no	Indirect (r)	no
DFEalpha: Keightley & Eyre-Walker. **2007**. Genetics. / Schneider et al. **2011**. Genetics.	yes	"yes"	Direct (simple)	no
Prfreq: Boyko et al. **2008**. Plos Genetics.	yes	no	Direct (complex)	no
δaδi: Gutenkunst et al. **2009**. Plos Genetics / Huber et al. **2017**. PNAS /Kim et al. **2017**. Genetics	yes	"yes"	Direct (complex)	"yes"
Grapes: Galtier. **2016**. Plos Genetics.	yes	yes	Indirect (r)	no
polyDFE: Tataru et al. **2017**. Genetics / Tataru & Bataillon. **2019**. Bioinformatics	yes	yes	Indirect (r)	yes
Anavar: Barton & Zeng. **2018**. MBE	yes	yes	Indirect (r)	no
DEFinitely. Lyn Fortier et al. **2019**. BioRxiv	yes	yes	Direct (complex)	yes

polyDFE



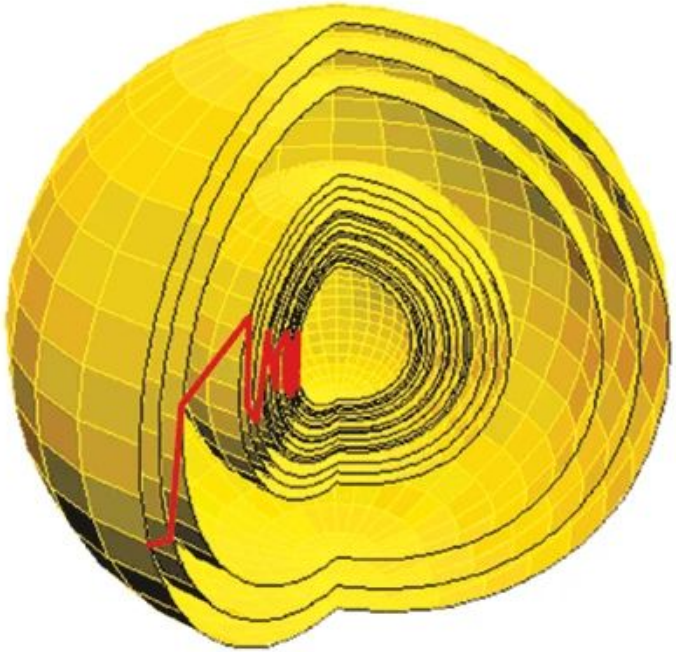
polyDFE, Tataru *et al.* 2016

polyDFE



polyDFE, Tataru *et al.* 2016

Previous DFE comparisons: Huber *et al.* 2017 and the FGM



Orr 2005

Previous DFE comparisons: Huber *et al.* 2017 and the FGM

Spp: humans, *Drosophila*, yeast, and mice.

Fisher's geometrical model (FGM) predictions fit well with real observations:

- Mutations in more complex organisms are on average more deleterious, since mutations are more likely to disrupt something important in a complex organism than in a simple one.

Previous DFE comparisons: Huber *et al.* 2017 and the FGM

Spp: humans, *Drosophila*, yeast, and mice.

Fisher's geometrical model (FGM) predictions fit well with real observations:

- Mutations in more complex organisms are on average more deleterious, since mutations are more likely to disrupt something important in a complex organism than in a simple one.
- More pleiotropic mutations will show smaller variation in s_d. Or, less pleiotropic mutations tend to be either close to neutral or very deleterious.

Previous DFE comparisons: Huber *et al.* 2017 and the FGM

Spp: humans, *Drosophila*, yeast, and mice.

Fisher's geometrical model (FGM) predictions fit well with real observations:

- Mutations in more complex organisms are on average more deleterious, since mutations are more likely to disrupt something important in a complex organism than in a simple one.
- More pleiotropic mutations will show smaller variation in s_d. Or, less pleiotropic mutations tend to be either close to neutral or very deleterious.
- Smaller populations will have a larger proportion of new beneficial mutations due to drift load.

Previous DFE comparisons: Huber *et al.* 2017 and the FGM

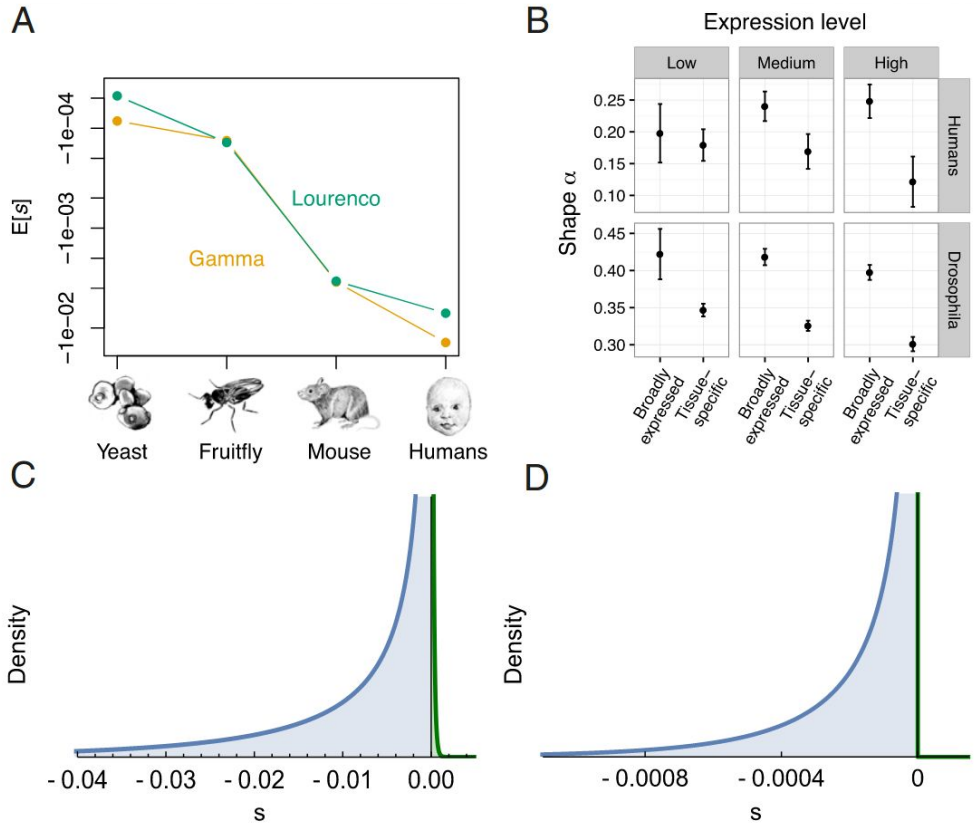
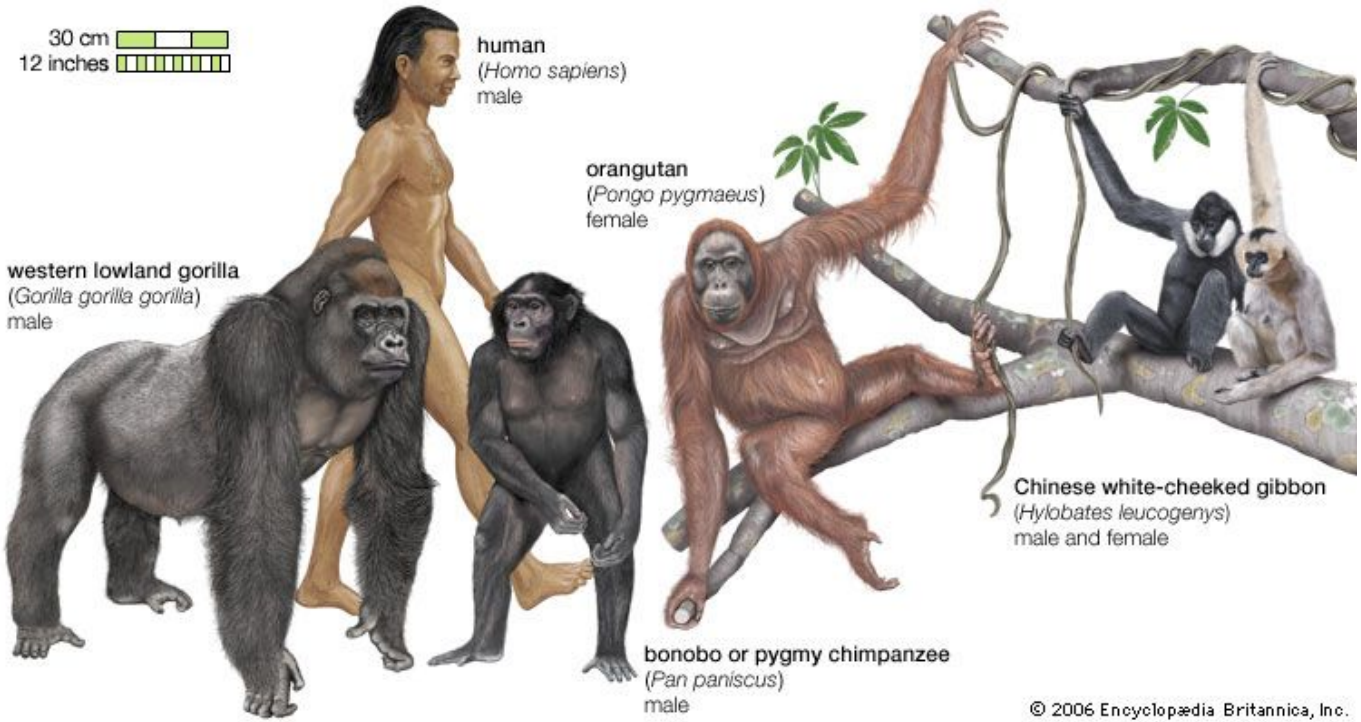


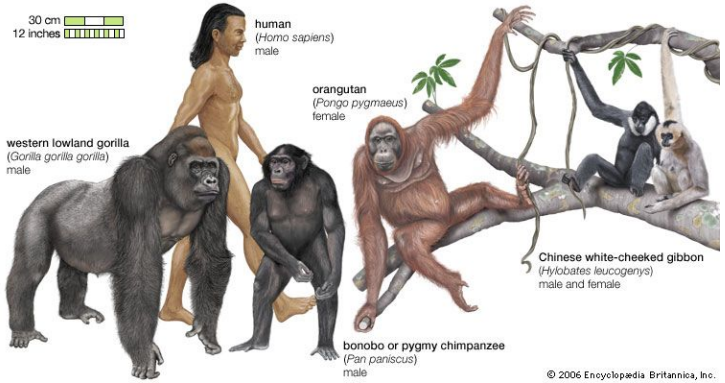
Fig. 4. Empirical support for FGM. (*A*) Both under the gamma DFE and the Lourenço et al. DFE, estimated average deleteriousness of mutations increases as a function of organismal complexity. (*B*) The shape parameter of the gamma DFE depends on the breadth of gene expression. Tissue-specific genes have a smaller shape parameter (α) than broadly expressed genes, supporting FGM. This pattern is consistent across overall expression levels. (*C* and *D*) By fitting the DFE of Lourenço et al., we can model slightly beneficial mutations in the DFE (green) that are thought to compensate for fixed deleterious mutations in species with small population size. We find support for a larger proportion of slightly beneficial mutations in the DFE of (*C*) humans than in (*D*) *Drosophila*.

Other DFE comparisons: Lyn Fortier *et al.* 2019

Great Apes are Great!

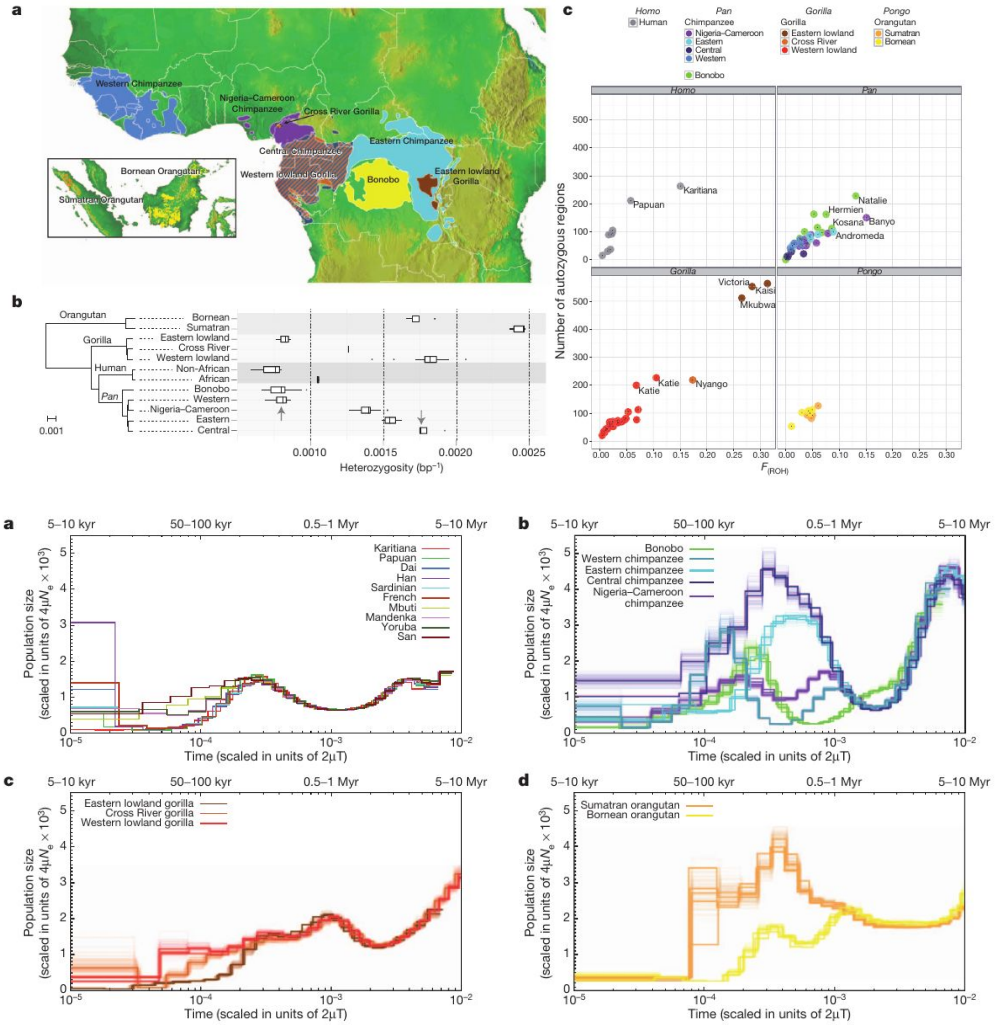


Great Apes are Great!



Data

Population genomic data from ~ 9,000 coding genes orthologous 1-to-1 across great apes, gibbons and macaques (Prado-Martinez *et al.* 2013). Eight chromosomes per population, nine populations.



Aim of the current study

Does $\mathcal{N}_e$ affect the full DFE all else being equal?

Aim of the current study

Does $\mathcal{N}_e$ affect the full DFE all else being equal?

1. Mutation rate
2. Recombination rate
3. Gene density
4. Gene number
5. Number of protein-protein interactions
6. Number of cell types

Aim of the current study

1. Is the shape of the deleterious DFE similar across great apes (FGM)?
2. Does the mean effect size of deleterious mutations ($S_d = 2N_e s_d$) scale proportionally to N_e (NNT)?
3. Is the rate of new beneficial mutations higher in low N_e populations (FGM)?
4. Is the effect of new beneficial mutations ($S_b = 2N_e s_b$) higher in large N_e populations (NNT)?

Mat&Met: How is the DFE estimated & compared?

Model	S_d	b	S_b	p_b	r_i	ε_{anc}	θ
2I	independent	independent	independent	independent	independent	independent	independent
2S	independent	shared	independent	independent	independent	independent	independent
3I	independent	independent	0	0	independent	0	independent
3S	independent	shared	0	0	independent	0	independent
3SS	shared	shared	0	0	independent	0	independent

Deleterious DFE (gamma)

Beneficial DFE (exp & discrete)

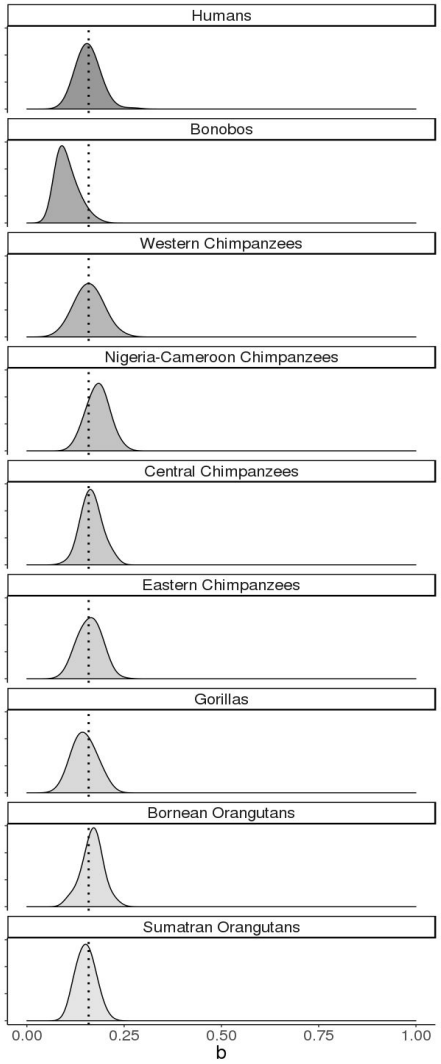
SFS distorters (demography, linkage, etc)

Population mutation rate ($4N_e\mu$)

polyDFEv2.0, Tataru & Bataillon 2019

Results

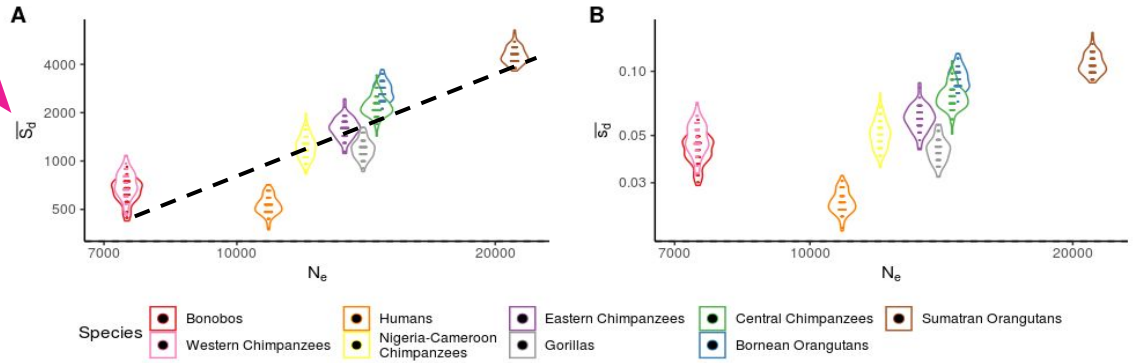
1. The shape of the deleterious DFE is stable across great apes. This is expected given the similar level of biological complexity across great apes.



Results

2. The population scaled mean effect size (S_d) of new deleterious mutations increases with the effective population size (N_e) as expected by the Nearly Neutral Theory (Ohta 1973).

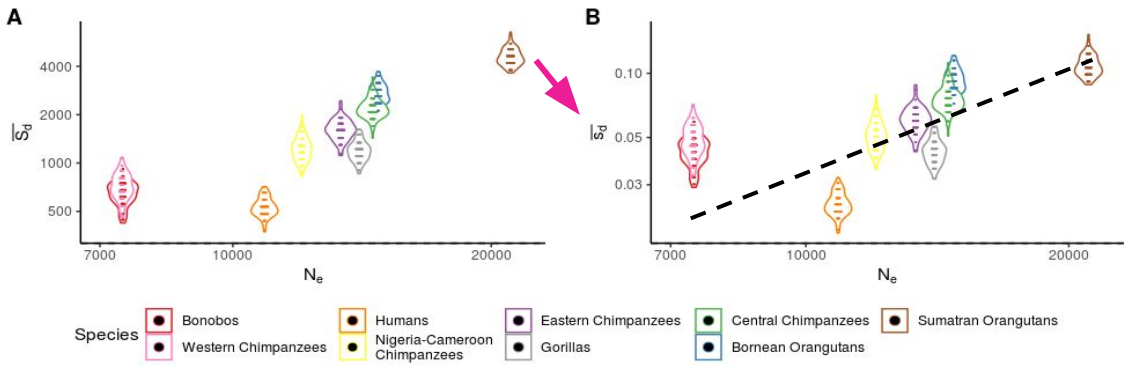
$$S_d = 2N_e s_d$$



Results

3. The mean effect size (s_d) of new deleterious mutations increases with the effective population size (N_e). This is unexpected!! Might be driven by positive epistasis.

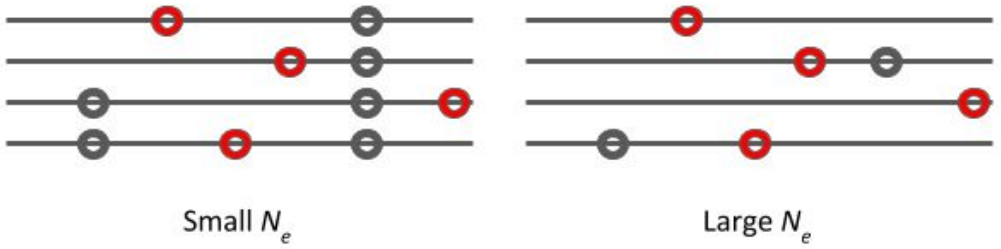
$$S_d = 2N_e s_d$$

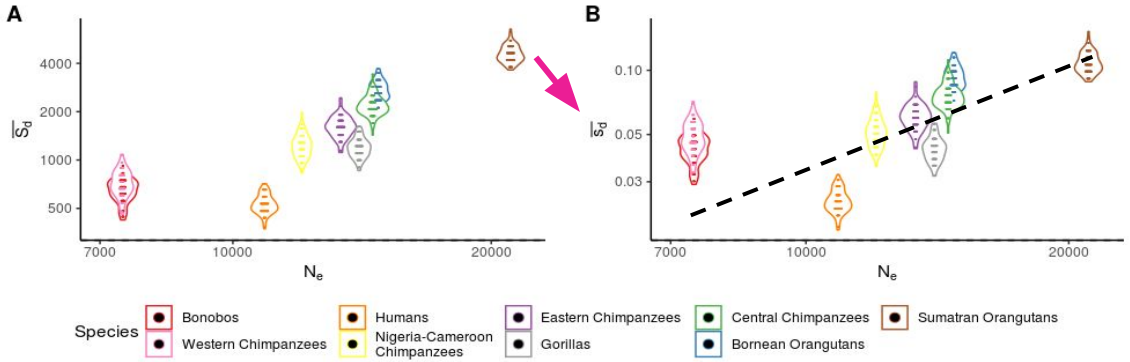


Results

3. The mean effect size (s_d) of new deleterious mutations increases with the effective population size (N_e). This is unexpected!! Might be driven by positive epistasis.

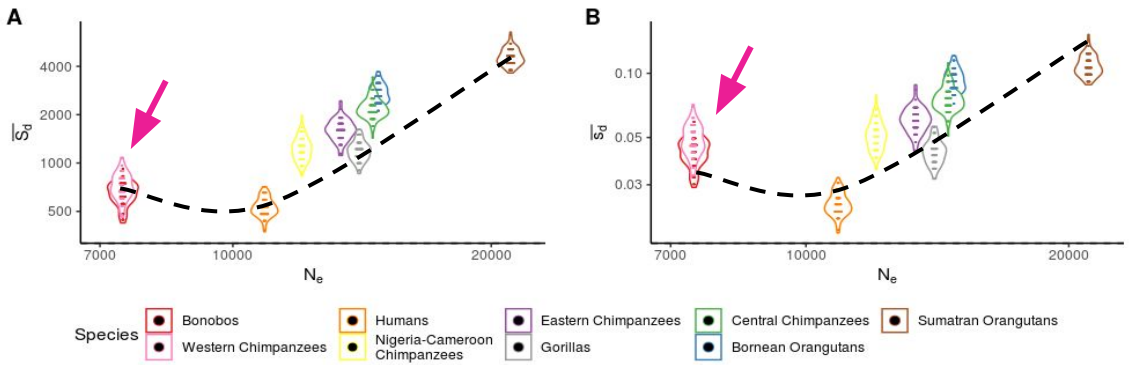
$$S_d = 2N_e s_d$$





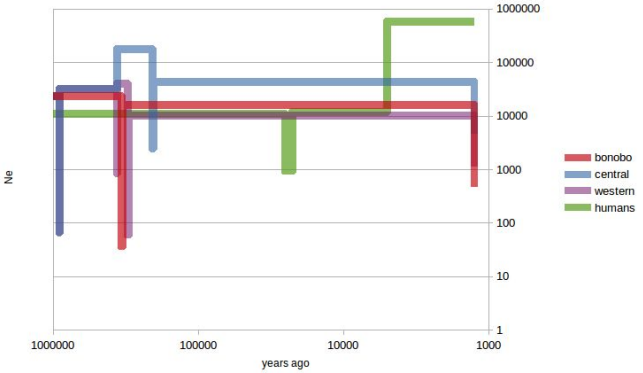
Results

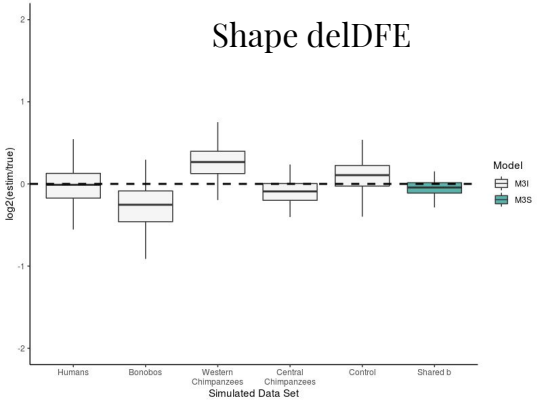
4. The strength of purifying selection increases in the smallest great apes. This might be due to the efficient purging of strongly deleterious recessive mutations in small populations.

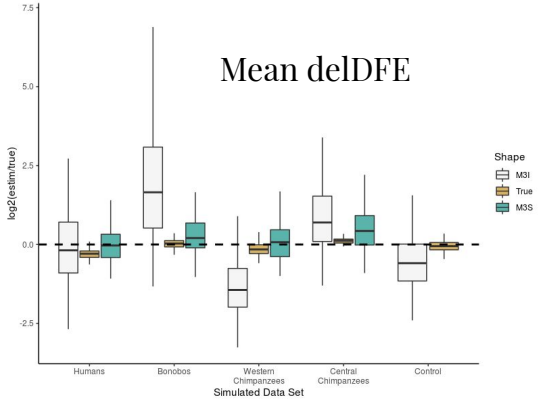


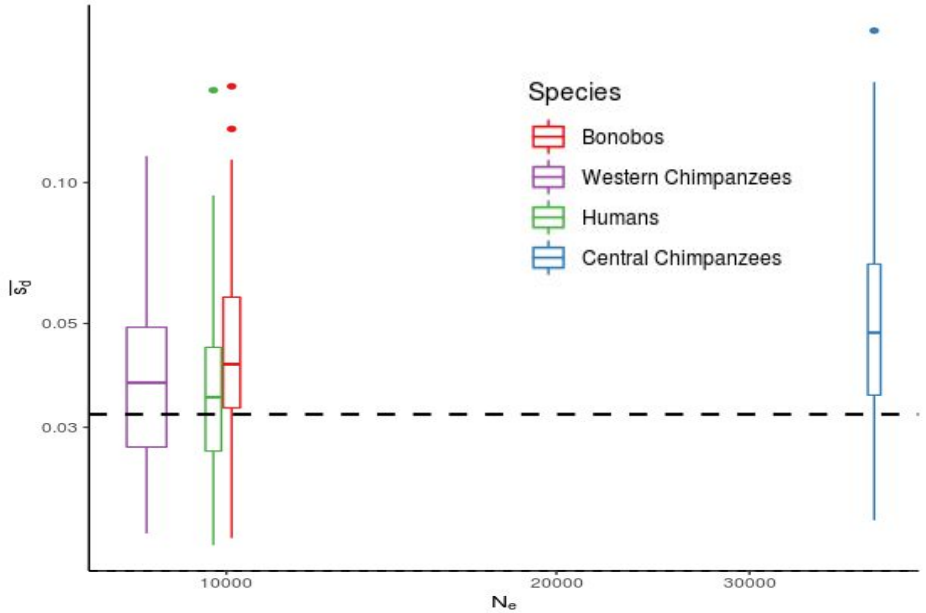
Results - Forward Simulations

de Manuel et al. 2016; Kim et al 2017.







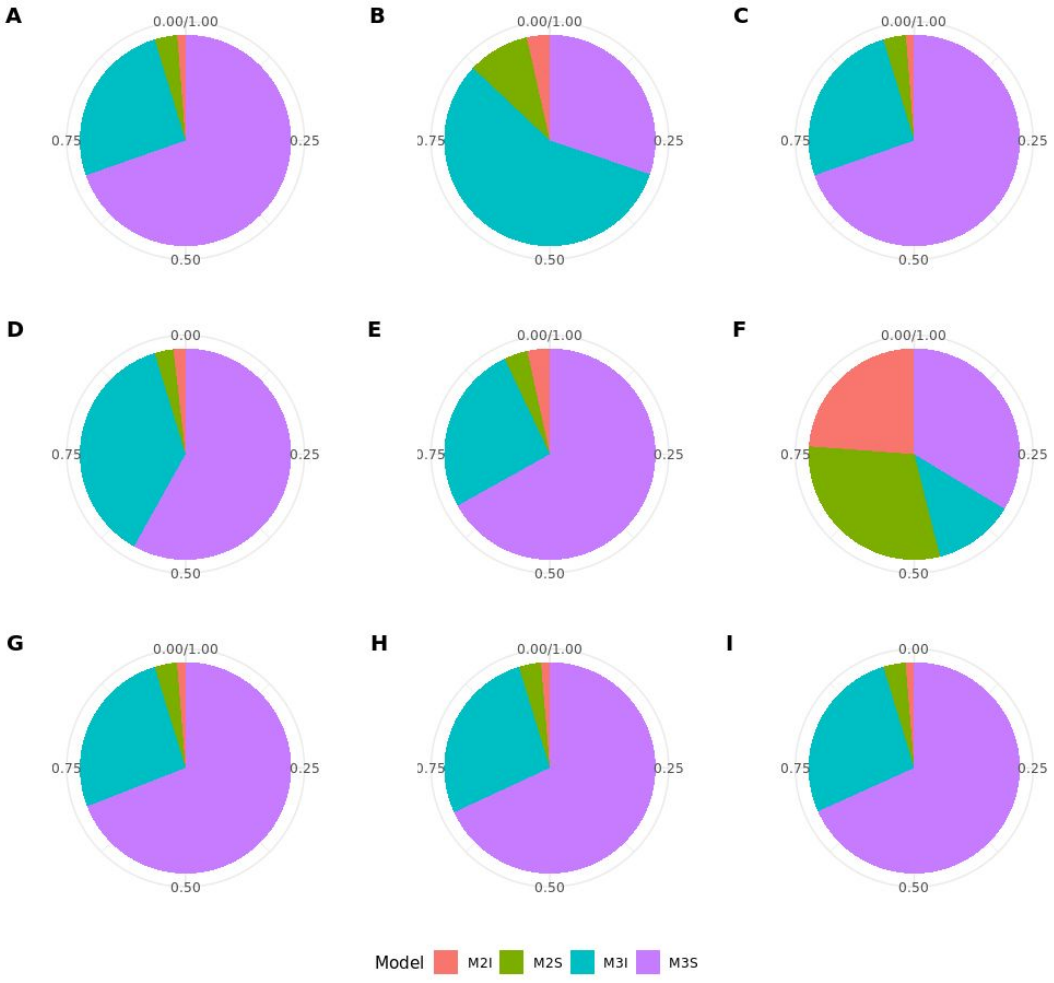


Results

5. 1-2% of new mutations are mildly beneficial in bonobos.

$$x_{avg} = \frac{\sum_j x_j e^{-1/2\Delta AIC_j}}{\sum_j e^{-1/2\Delta AIC_j}}$$

$$\Delta AIC_j = AIC_j - min_j(AIC_j)$$

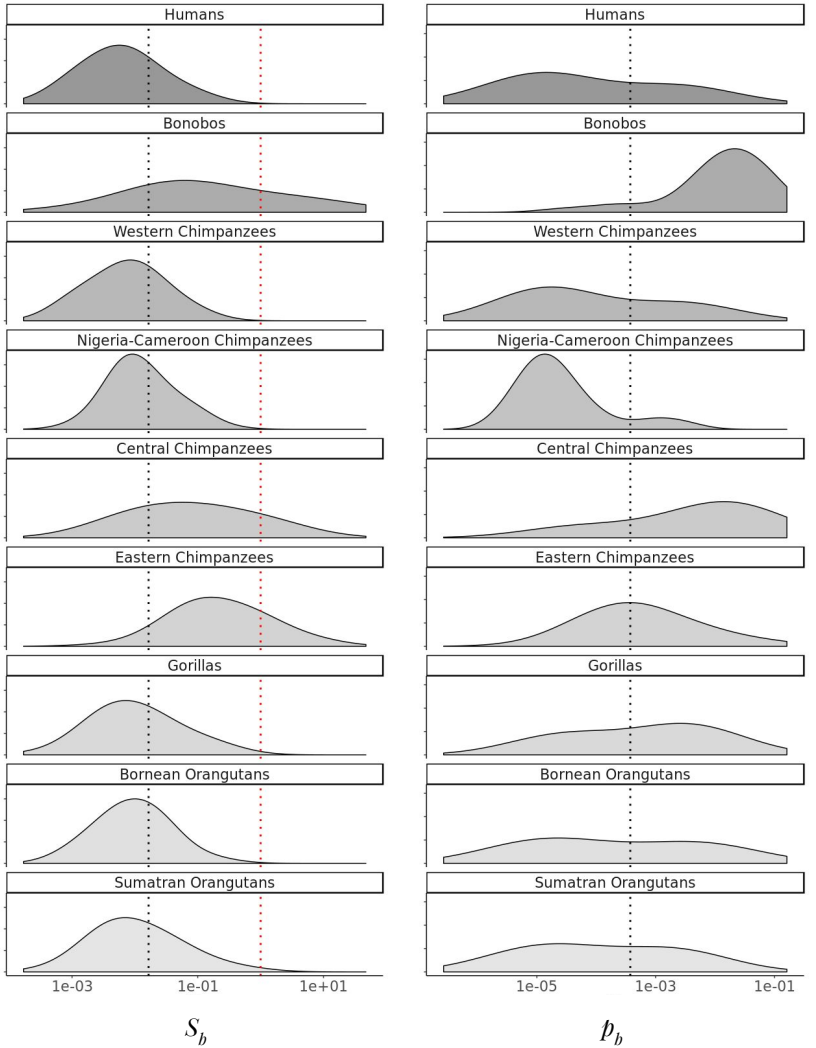


Results

5. 1-2% of new mutations are mildly beneficial in bonobos.

$$x_{avg} = \frac{\sum_j x_j e^{-1/2\Delta AIC_j}}{\sum_j e^{-1/2\Delta AIC_j}}$$

$$\Delta AIC_j = AIC_j - min_j(AIC_j)$$



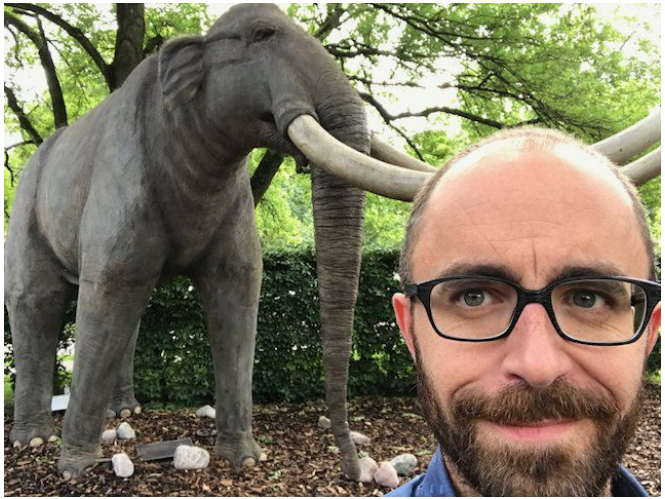
Aim of the current study

1. Is the shape of the deleterious DFE similar across great apes (FGM)? **YES, consistent with these species having similar level of biological complexity.**
2. Does the mean effect size of deleterious mutations ($S_d = 2N_e s_d$) scale proportionally to N_e (NNT)? **NO, positive epistasis also contributes.**
3. Is the rate of new beneficial mutations higher in low N_e populations (FGM)? **Probably but better answer with larger samples.**
4. Is the effect of new beneficial mutations ($S_b = 2N_e s_b$) higher in large N_e populations (NNT)? **We can not answer this question yet. Larger samples.**

Conclusions

Our study demonstrates the simple, but perhaps underappreciated fact that mutational effects are very dynamic even between closely related species.

Collaborators and funding



Kasper Munch

Aarhus University (Denmark)

Paula Tataru

Moises Coll-Macia

Thomas Bataillon

This work was supported by the Danish Council For Independent Research (grant number 4181-00358).

Thanks!