Title : Is adaptation limited by mutation? A timescale dependent effect of genetic diversity on the adaptive substitution rate in animals.

Authors : Rousselle M[1], Simion P[1,2], Tilak MK[1], Figuet E[1], Nabholz B[1], Galtier N[1].

[1]ISEM, Univ. Montpellier, CNRS, EPHE, IRD, Montpellier, France.

[2]LEGE, Department of Biology, University of Namur, 5000 Namur, Belgium.

Corresponding author: Marjolaine Rousselle

marjolaine.rousselle@umontpellier.fr

ABSTRACT

Whether adaptation is limited by the supply of beneficial mutations is a long-standing question of evolutionary genetics, which more generally relates to the determination of the adaptive substitution rate and its relationship with the effective population size N_e. The empirical evidence so far is equivocal, with some but not all studies supporting a higher adaptive substitution rate in large-N_e than in small-N_e species.

We gathered coding sequence polymorphism data and estimated the adaptive amino-acid substitution rate ω_a, in 50 species from ten distant groups of animals with markedly different population mutation rate θ. We reveal the existence of a complex, timescale dependent relationship between species adaptive substitution rate and genetic diversity. On the one hand, we report a positive relationship between ω_a and θ among closely related species, indicating that adaptation is indeed limited by the supply of beneficial mutations, but this was only true in relatively low-θ taxa. On the other hand, we uncovered a weak negative correlation between ω_a and θ at a larger taxonomic scale. This result is consistent with predictions of Fisher's geometrical model and suggests that the proportion of mutations that are beneficial scales negatively with species long-term N_e.

Key words: adaptive substitution rate, beneficial mutations, effective population size, Mc-Donald and Kreitman, animals.

INTRODUCTION

It is widely accepted that adaptation is more efficient in large populations. Firstly, because they produce a greater number of mutants per generation, large populations are more likely than small ones to newly find the required alleles, if missing from the gene pool. Secondly, because they tend to be genetically more diverse, large populations are more likely to carry the alleles needed to respond to environmental changes (1). Lastly, the fixation probability of beneficial mutations is higher in large than in small populations due to a weaker effect of genetic drift in the former. So, be it from standing variation or from *de novo* mutations, one would intuitively expect to observe a higher rate of accumulation of adaptive changes, on average, in large than in small populations (2).

Under a simple population genetic model, in a population of effective size N_e, mutations of selection coefficient $s >> 1/N_e$ should accumulate at rate $\sim 4N_e\mu_a s$, where μ_a is the adaptive mutation rate – i.e., the adaptive substitution rate should scale linearly with $N_e\mu$ (where μ is the total mutation rate) (3).

This rationale implicitly assumes that the rate of adaptation is limited by the supply of new mutations, i.e., the population mutation rate $\theta=4N_e\mu$ (4). It might be, however, that the amount of genetic diversity available in all or most existing populations is sufficient for adaptation, and/or that the ability to adapt to environmental changes is determined in the first place by factors independent from the effective population size, such as the magnitude or frequency of perturbations, the finite set of possible genotypes an organism can reach, or the ability of populations to combine favorable alleles across loci via recombination (5–10). Finally, it has been suggested that the rate of adaptive mutation, μ_a, might be negatively correlated with N_e, which further complicates the situation. This is because small populations tend to accumulate deleterious mutations, and the resulting load could offer the opportunity for adaptive, compensatory mutations to appear and spread irrespective of environmental perturbations (9). Theoretical models, therefore, can predict either a positive, a negative, or a lack of relationship between population size and adaptive rate, depending on the underlying assumptions.

Molecular data offer an unique opportunity to empirically evaluate the correlation between the rate of adaptation and θ, and thus to test whether adaptation is indeed limited by mutation. More efficient adaptation in large populations should manifest itself by an increased rate of protein evolution, which can be estimated from coding sequence alignments. The ratio of non-synonymous (i.e. amino-acid changing, dN) to synonymous (i.e. amino-acid conservative, dS) substitution rates, often called ω, is a measure of the protein evolutionary rate that controls for the effects of divergence time and mutation rate. ω, however, is not only influenced by adaptation, but also by the strength and efficiency of purifying selection against deleterious alleles. To account for this, McDonald and Kreitman (1991, MK) (10) suggested including within-species polymorphism in the analysis. Because adaptive mutations are rare and not particularly long-lived, they are expected to contribute negligibly to the pool of segregating alleles. The ratio of non-synonymous to synonymous polymorphism, therefore, provides an estimate of the expected ω under near-neutrality, i.e., in absence of adaptation, called ω_{na} (for non-adaptive). Subtracting the nearly-neutral

expectation ω_{na} from the observed ω provides an estimate of the adaptive rate, ω_a, and the proportion of adaptive substitution rate, α (11).

Subsequent improvements of the MK method were intended to account for a number of factors potentially confounding the estimation of ω_{na}, including the prevalence of slightly deleterious segregating alleles and recent demographic effects (14–21). Improved methods explicitly model the distribution of fitness effect (DFE) of non-synonymous mutations, and take information not only from the number of synonymous and non-synonymous single nucleotide polymorphisms (SNPs), but also from the distribution of allele frequencies across SNPs – the so-called site frequency spectra (SFS). The ω_a statistics has a high sampling variance (22) and its estimation can be biased by various factors, such as fluctuating population size (12,23,24) and GC-biased gene conversion (25–27). In particular, one key assumption of the MK approach is that the long-term N_e, which determines ω, is equal to the short-term N_e and can therefore be estimated from polymorphism data. This appears unlikely to be generally true, and ancient fluctuations in N_e could in principle fault the MK rationale (12,23,24). Eyre-Walker (24) theoretically considered the problem of a single ancient change in N_e and showed that an expansion in population size, even if old, could lead to overestimation of the adaptive substitution rate, a bias that could create spurious positive correlation between ω_a and N_e and that we need to keep in mind when interpreting this type of estimates.

The first applications of the MK method to large-scale data sets indicated that the adaptive rate is higher in *Drosophila* than in humans (11,13,14), consistent with the prediction of more efficient adaptation in large populations and with the hypothesis that mutation limits adaptation. These studies, however, were focusing on the $\alpha=\omega_a/(\omega_a+\omega_{na})$ statistics, i.e., the proportion of amino-acid substitutions that result from adaptation. α, however, is influenced by ω_{na} as well as ω_a, and a lower α in humans than in *Drosophila* might mainly reflect a higher rate of non-adaptive amino-acid substitution in the former. Indeed, purifying selection against deleterious mutations is expected to be less effective in small populations due to increased genetic drift (28). Comparative studies focusing on ω_a have only revealed tenuous positive effects of θ on the adaptive rate in mammals, flies and plants (29–31). The largest analysis of this sort used 44 pairs of non-model species of animals occupying a wide range of θ (18). This study reported a significantly positive relationship between θ-related life history traits and α, consistent with previous literature, but this was entirely due to the non-adaptive component. Galtier (18) failed to detect any effect of θ on ω_a, despite using various models for the distribution of fitness effects and accounting for a number of potential

confounding factors - a result that does not support the hypothesis that adaptation is limited by mutation. Of note, the species sampled in this study belong to divergent families and phyla of animals. They differ much in terms of ecology and genome architecture, which might confound the relationship between ω_a and θ (18).

So, the evidence so far regarding the relationship between adaptive rate and population mutation rate is equivocal. Existing comparative studies have adopted distinctive methodological approaches, both in terms of species sampling and estimates of the adaptive rate. In particular, these studies have been conducted at different evolutionary scales, which might contribute to explain their somewhat discordant results. In the short term, an increase in N_e is expected to boost the adaptive substitution rate if the supply of new mutations is limiting. In the long run, differences in N_e could also lead to changes in the DFE, and particularly the proportion of beneficial mutations, due to small-N_e species being pulled away from their fitness optimum via genetic drift (11,18,32). How these two opposing forces interact and combine to determine the relationship between ω_a and θ is still unknown, in absence of a multi-scale study.

In this study, we propose to test the effect of evolutionary scale on the relationship between the adaptive substitution rate (ω_a) and the population mutation rate (θ). For this, we gathered coding sequence polymorphism data in 4 to 6 species from each of ten distant groups of animals with markedly different θ. Our results reveal that the relationship between ω_a and θ varies depending on the considered taxonomic scale, i.e. depending on whether we compare closely related species or distantly related taxa. We report a positive relationship between ω_a and θ within groups, the strength of which weakens as θ increases, indicating that adaptation is limited by beneficial mutations in small-θ species of animals. At a larger taxonomic scale, in contrast, we report a weak negative correlation between ω_a and θ, with primates and ants showing a higher adaptive substitution rate than mussels and fruit flies, for instance. This is consistent with the hypothesis that long-term N_e influences the DFE, and particularly the proportion of adaptive mutations.

RESULTS

1. Data sets

We assembled a data set of coding sequence polymorphism in 50 species from ten taxonomic groups, each group including 4 to 6 closely-related species (**Table S1**). The ten taxa we analyzed are Catharrhni (Mammalia, hereafter called “primates”), Passeriformes (Aves, hereafter called “passerines”), Galloanserae (Aves, hereafter called “fowls”), Muroidea (Mammalia, hereafter called “rodents”), Lumbricidae (Annelida, hereafter called “earth worms”), *Lineus* (Nemertea, hereafter called “ribbon worms”), *Mytilus* (Mollusca, hereafter called “mussels”), Satyrini (Lepidoptera, hereafter called “butterflies”), *Formica* (Hymenoptera, hereafter called “ants”), and *Drosophila* (hereafter called “flies”).

Data for five groups (primates, passerines, fowls, rodents and flies) were obtained from public databases. Data for the other five groups were newly generated via exon capture in a total of 242 individuals from 22 species (**Table 1**) and we obtained sufficient data for 216 of them (~89%). The average coverage was of 9X in ants, 23X in butterflies, 10X in earth worms, 28X in ribbon worms and 26X in mussels (average of median coverage per species). The percentage of targeted coding sequences for which at least one contig was recovered varied from 31.9% to 88.2% across species (median=78.8%, **Table 1**).

Species	Group	Targeted transcripts	Recovered transcripts	Percentage of recovered among targeted transcripts
Formica fusca	ants	1810	1427	78.8
Formica sanguinea	ants	1810	1396	77.1
Formica pratensis	ants	1810	1398	77.2
Formica cunicularia	ants	1810	1406	77.7
Maniola jurtina	butterflies	2235	1921	86.0
Melanargia galathea	butterflies	2235	1713	76.6
Pyronia tithonus	butterflies	2235	1823	81.6
Pyronia bathseba	butterflies	2235	1864	83.4
Aphantopus hyperanthus	butterflies	2235	1772	79.3
Allolobophora chlorotica L1	earth worms	2955	2293	77.6
Allolobophora chlorotica L2	earth worms	2955	2315	78.3
Allolobophora chlorotica L4	earth worms	2955	1732	58.6
Aporrectodea icterica	earth worms	2955	2321	78.5
Lumbricus terrestris	earth worms	2955	943	31.9
Lineus sanguineus	ribbon worms	1725	1251	72.5
Lineus ruber	ribbon worms	1725	1521	88.2
Lineus lacteus	ribbon worms	1725	1516	87.9
Lineus longissimus	ribbon worms	1725	1505	87.2
Mytilus galloprovincialis	mussels	2181	1820	83.4
Mytilus edulis	mussels	2181	1721	78.9
Mytilus trossulus	mussels	2181	1740	79.8
Mytilus californianus	mussels	2181	1808	82.9

Table 1: Summary of the number of targeted transcripts recovered in the capture experiment.

We assessed contamination between samples from distinct species using the software CroCo (33). Overall, inter-groups connections in **Figure S1** indicates a low level of cross-contamination: when there are connections between taxonomic groups, they concern on average 38 contigs identified as contaminants, the worst case being the 172 contigs identified as contaminants between the assembly of *Lineus sanguineus* and *Mytilus galloprovincialis*. The connections between assemblies from closely related species are very likely to be false positive cases, especially since the intensity of the within-group connections is congruent with the phylogenetic distance between species within taxa.

Regardless, all the contigs identified as potential contaminants were excluded from the dataset in downstream analyzes by safety.

Within each group, we focused on orthologous contigs (**Table S2**), predicted open reading frames, ands called the diploid genotypes of individuals ar every coding position. SNPs counts obtained after genotyping are summed up in **Table S3**. Only in two species did we obtain less than a thousand SNPs, the minimum being 153 for *Lineus longissimus*, for which we could recover data for only six individuals.

We recovered an average of 8,459 SNPs per species in ants, 7,950 in butterflies, 4,763 in earth worms, 8,347 in ribbon worms, 19,750 in mussels, 10,191 in primates, 25,534 in rodents, 40,870 in passerines, 8,488 in fowls and 195,398 in flies.

In conclusion, the capture experiment seems to be suitable to recover population coding sequences data for several closely related species - here, the maximum divergence between species within a taxonomic group was 0.2 subst./site, i.e. the divergence between *Lumbricus terrestris* and *Allolobophora chlorotica L1*.

2. Between-groups relationship between the population mutation rate (θ) and the adaptive substitution rate (ω_a)

Two strategies were adopted to combine SFS information from distinct species in a group-level estimate of ω_a, thus accounting for the problem of phylogenetic non-independence between species. For both strategies, we first calculated the dN/dS ratio ω at the group-level, i.e., averaging across all branches of the tree (see Material and Methods). The first one, which we call $\omega_{a[P]}$, was obtained by pooling the SFS from distinct species within a group, separately for synonymous and non-synonymous SNPs (as in (34)), before fitting the model and estimating the parameters. We then computed the relationship between $\omega_{a[P]}$ estimates and the across-species average nucleotide diversity, π_s, taken as an estimate of θ. We detected a significant negative relationship between $\omega_{a[P]}$ and the across-species average nucleotide diversity, π_s, taken as an estimate of θ (regression test, r^2=0.4, p-value=0.029) (**Figure 1A**).

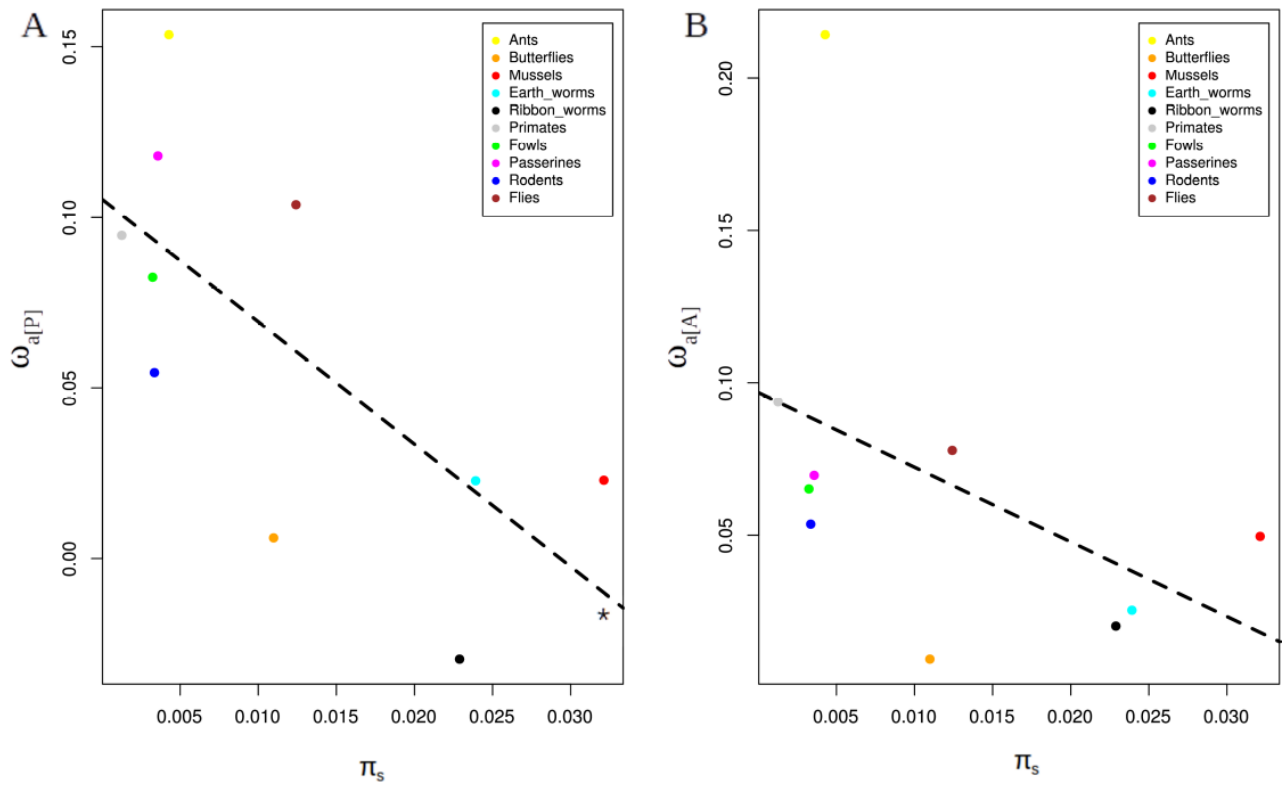

Figure 1: Relationship between group-level ω_a and group-level π_s.

A: ω_a is estimated by pooling SFS across species within a group ($\omega_{a[P]}$).

B: ω_a is estimated via the averaging of ω_{na} across species within a group ($\omega_{a[A]}$).

Group level π_s is estimated by averaging species-level π_s across closely related species. The dotted lines represent the regression across taxonomic groups and * symbols indicate when it is significant.

Recent studies focusing on birds and more recently on primates indicate that GC-biased gene conversion (gBGC) may lead to an overestimation (25,26) or an underestimation of ω_a (27). Interestingly, gBGC does not affect genomic evolution with the same intensity in all organisms (35). To avoid biases in the estimation in species where gBGC is active, we restricted the SNPs and substitution data to GC-conservative changes, which are not influenced by gBGC. Here again, the correlations, even if not significant, were in line with a negative relationship between $\omega_{a[P]GC\text{-}conservative}$ and proxies of the long-term N_e (**Figure S2**).

Our second estimate of the adaptive rate at group-level, which we called $\omega_{a[A]}$, was obtained by calculating the across-species arithmetic mean of ω_{na} within a group, then subtracting this average from ω. It can be shown that $\omega_{a[A]}$ is an unbiased estimate of the adaptive rate with fluctuating

population size if the pace of fluctuations is sufficiently slow (Supplementary Material **Box S1**). We found a non significant negative correlation between $\omega_{a[A]}$ and $\omega_{a[A]GC\text{-}conservative}$ and π_s (**Figure 1B and Figure S3**). Overall, the between-group analysis confirmed the absence of a positive relationship between ω_a and θ at the between-phyla scale in animals, and even suggested the existence of a weak, negative relationship.

3. Within-group relationship between θ and ω_a

To assess the within-group effect of π_s on ω_a, we performed an analysis of covariance (ANCOVA) with taxonomic group as a categorical independent variable, as in (29). The principle of this analysis is to fit a set of parallel lines (one for each taxonomic group) and test whether their common slope is significantly different from zero. Additionally, we tested if the relationship between ω_a and π_s or life history traits differs between taxonomic groups by testing whether the lines have different intercepts.

Using this strategy, we found that ω_a and both π_s and $\log_{10}(\pi_s)$ are significantly positively correlated when using only GC-conservative mutations (ANCOVA p-value=0.028 and 0.0031, respectively) (**Figure 2**). ω_a is only marginally positively correlated to $\log_{10}(\pi_s)$ when using all mutations (ANCOVA p-value=0.076). We also found that there is a significant variation between the intercepts (ANCOVA p-value<0.001), as well as a significant interaction between the dependent variable and categorical independent variable (ANOVA p-value=0.016). Those results support the existence of a positive relationship between ω_a and θ within groups, with the slope of the relationship differing between groups. This is consistent with the hypothesis that within a group, higher-θ species are more likely to find and fix adaptive substitutions than low-θ species, in line with the hypothesis that mutation limits adaptation. **Figure 2** shows that the slopes of the within-group ω_a/θ correlations decrease with group-level π_s, and we actually find a significant negative correlation between these two quantities (Spearman correlation coefficient=-0.77 p-value=0.014). This interestingly suggests that the limitation of adaptation by the income of adaptive mutations is effective and strong in small-θ groups (e.g. primates, rodents, ants), but not in high- groups of animals (e.g. flies, mussels, butterflies), where the ω_a/θ relationship is essentially flat. (**Figure 2**).

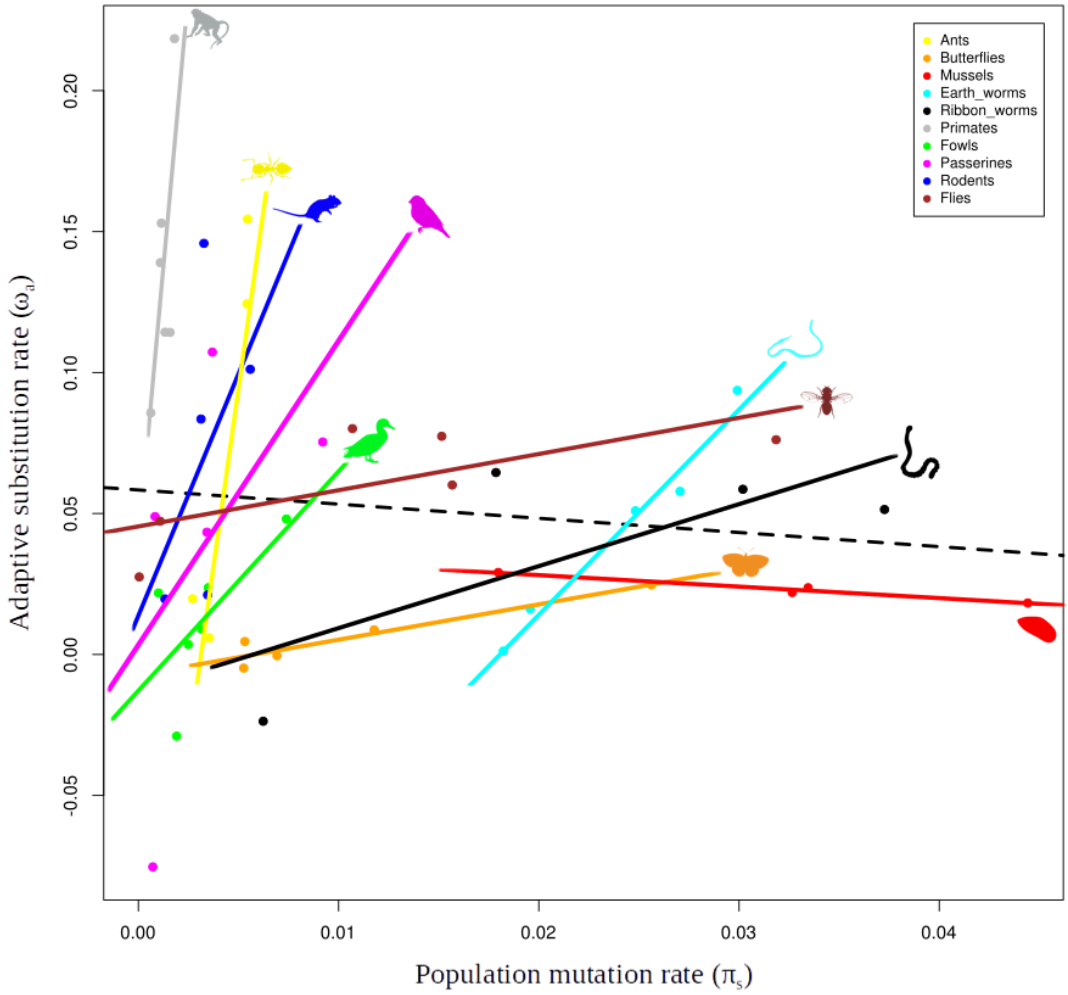

Figure 2: Relationship between species-level ω_a and π_s.

ω_a is estimated using only GC-conservative mutations. The dotted line represents the regression across all species, and full lines represent the regression within each taxonomic groups.

When analyzing the per-species non-adaptive substitution rate in the same manner, we found a global negative relationship between ω_{na} (using both all mutations and only GC-conservative mutations) and π_s (regression test r^2=0.33, p-value=7.7e-6), and also a significantly negative relationship within groups (ANCOVA p-value=0.0018) (**Figure S4**). This is consistent with the expectations of the nearly neutral theory of evolution (28), and with previous empirical results (18,36). The estimated ratio of adaptive to total non-synonymous substitutions, α, behaved more or less similarly to ω_a (**Figure S5**).

4. Control for fluctuations in N_e

We were concerned that the correlation between ω_a and π_s might have arisen as an artifact generated by past fluctuations in population size (23,24). To test this, we simulated coding sequence evolution under several demographic scenarios with four regimes demographic fluctuations, with a three or thirty-fold ratio between the low and high N_e, and a high or low long-term N_e (see Material and Method and **Figure S6**). We found that the only scenario where demographic fluctuations could lead to a detectable positive correlation between ω_a and π_s is the one with the highest long-term N_e and highest difference between the low and high N_e (see **Figure S7 panel B,** regression test r^2=0.07, p-value=0.0095). The correlation disappeared when we used a ten-fold smaller long-term N_e, whereas we empirically observed that the correlation between ω_a and π_s is stronger for small long-term N_e groups (**Figure 2**). These simulations therefore suggest that ancient demographic fluctuations cannot explain our finding of a positive within-group correlation between ω_a and π_s in low-θ groups.

5. Relationship with life history traits

We used several life history traits known to be correlated with species long-term effective population size (37). In our data set, all life history traits are indeed correlated with π_s (Spearman correlation p-value, propagule size: 1.1e-12, adult size: 0.00043, longevity: 0.055, body mass: 0.0047, fecundity: 9.4e-6). When estimating per-group ω_a, we did not find any significant relationship with life history traits, but the sign of correlation coefficients were indicative of a negative relationship between the long-term N_e and both ω_a and $\omega_{a[GC\text{-}conservative]}$ (**Figure S2 and S3**). When considering all mutations and all 50 species, we found a negative relationship between ω_a and $\log_{10}$ transformed fecundity (regression test, r^2=0.094, p-value=0.038), as well as a positive relationship with $\log_{10}$ transformed longevity (regression test, r^2=0.10, p-value=0.022), $\log_{10}$ transformed propagule size (regression test, r^2=0.13, p-value=0.0073) (**Figure 3**). When using only GC-conservative mutations, the relationships were similar, while only significant with fecundity (regression test, r^2=0.11, p-value=0.0026) (**Figure 3**). When controlling for phylogenetic inertia, those relationships lost significance.

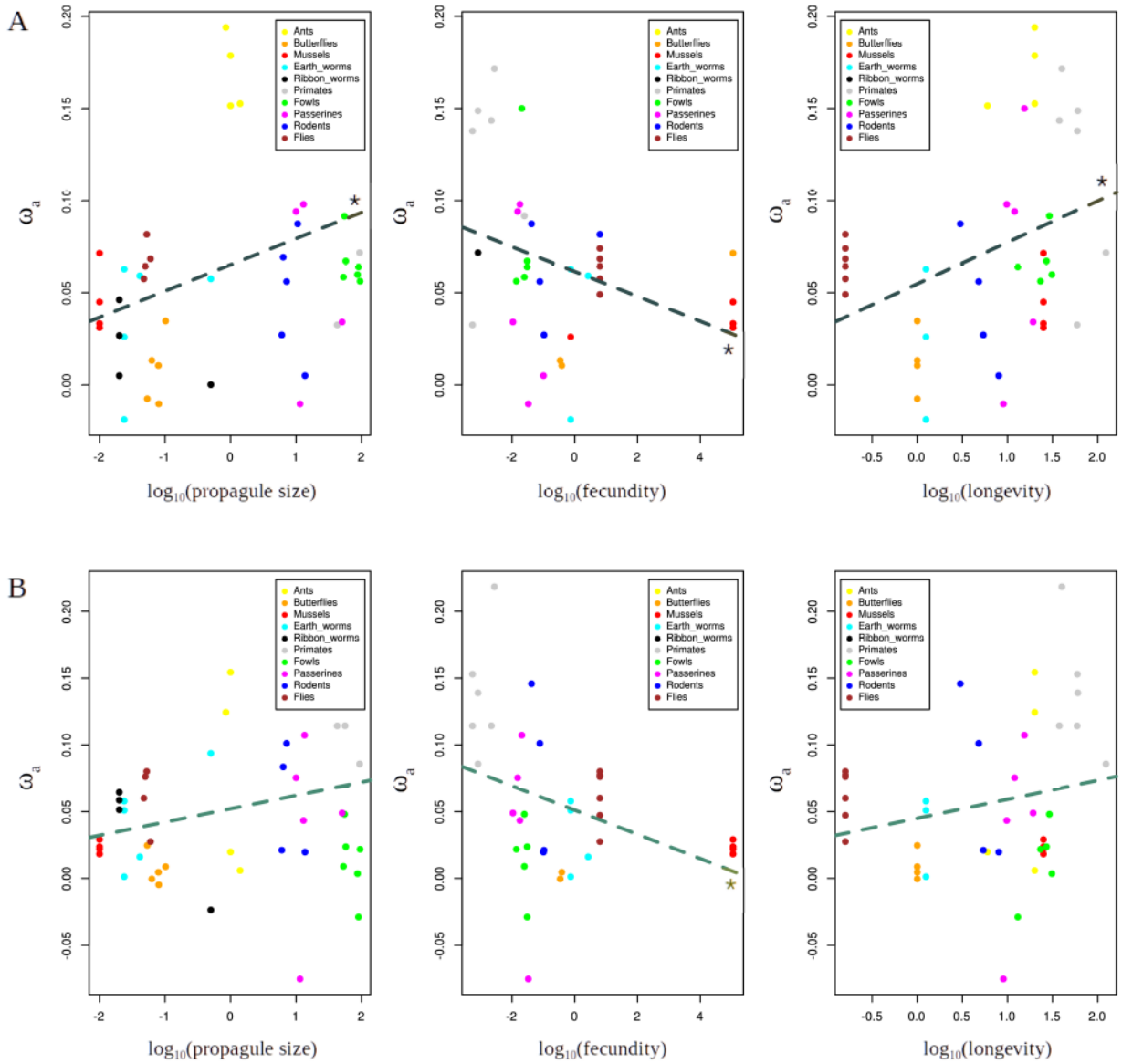

Figure 3: Relationship between species-level ω_a and life history traits.

A: ω_a is estimated using all mutations.

B: ω_a is estimated using only GC-conservative mutations.

The dotted lines represent the regression across all species and * symbols indicate when the regression is significant.

We also found a significant positive relationship between ω_{na} and fecundity (regression test r^2=0.31, p-value=0.00028), and a significant negative relationship between ω_{na} and propagule size (regression test r^2=0.13, p-value=0.0072) and body mass (regression test r^2=0.10, p-value=0.024), which was also true when using only GC-conservative mutations: significant negative relationships between $\omega_{na[GC\text{-}conservative]}$ and propagule size (regression test r^2=0.21, p-value=0.0006802), longevity

(regression test r^2=0.12, p-value=0.016), body mass (regression test r^2=0.08, p-value=0.045) and a significant positive relationship with fecundity (regression test r^2=0.36, p-value=6.236e-05) (**Figure S4**).

DISCUSSION

1. Influence of θ on ω_a : a two-scale mechanism

In this study, we analyzed a novel, 50-species population genomic data set to evaluate the relationship between the adaptive substitution rate and the population mutation rate and test the hypothesis that mutation limits adaptation in natural populations of animals. Using an exon capture strategy, we were able to generate large-scale synonymous and non-synonymous SFS data in 22 non-model species from five groups of non-vertebrates, which were added to five existing datasets in mammals, birds and flies. In all ten groups we started from raw data and *de novo* performed genome/transcriptome assembly, read mapping, SNP/genotype calling, and population genetic analyses.

We found that the relationship between ω_a and θ depends on the considered timescale. On the one hand, at a recent evolutionary scale (i.e., neutral divergence <0.2 subst./site), we found a significant positive correlation between ω_a and π_s (**Figure 2**). Interestingly, the slope of the relationship differed significantly among taxonomic groups, and this slope itself was negatively correlated with the group average π_s. On the other hand, estimations of the adaptive substitution rate at group level revealed a weak but consistently negative relationship between ω_a and π_s - and life history traits correlated with the long-term N_e (**Figure 1** and **3**). This time scale-dependent behavior of the ω_a/θ relationship is here demonstrated via the analysis of a single, multi-scale dataset, somehow reconciling earlier taxon-specific studies on the subject (4,8,18,29–31,38).

2. Relationship between θ and ω_a: a real causative link or an artifact ?

Using an ANCOVA analysis, we found that the slopes of the relationship between ω_a and π_s within each taxonomic group were significantly different from zero, demonstrating the existence of a positive link between ω_a and π_s within groups (**Figure 2**). We were concerned that this relationship may result from a bias of the MK approach, instead of a true biological signal. Indeed, the MK

approach implicitly assumes that the regime of selection/drift has been constant over the considered time period, i.e. since the divergence between the focal and outgroup species. However,if the selection/drift regime had changed (via a change in effective population size, for instance) between the period during which divergence was built and the period during which polymorphism was built, this could lead to overestimation or underestimation of ω_a (23,24). Here, we used the so-called r_i's nuisance parameters (39) to control for recent changes in N_e.

In contrast, ancient N_e changes that affect coding sequence divergence are virtually impossible to trace – an ancient bottleneck, for instance, is expected to result in an increased dN/dS ratio due to the accumulation of deleterious substitutions, similarly to an episode of positive selection. We showed in a previous study based on simulations that ancient demographic fluctuations can lead to severely overestimated α and ω_a, an upward bias which is exacerbated when the true adaptive substitution rate is low (23). Additionally, it has been shown by modeling single changes in N_e that in presence of slightly deleterious mutations, an increase in N_e in the past could yield spurious evidence for positive selection, which can lead to a spurious positive correlation between ω_a and π_s (24).

We used simulations to test if demographic fluctuations could lead to such a correlation. We simulated several scenarios of demographic fluctuations with different periods of fluctuations, two levels of severeness of fluctuations (two different ratios between low and high N_e phases) and two levels of long-term N_e. We found that the only scenario where a positive relationship between ω_a and π_s is detected is the one with the highest long-term N_e and the highest difference between low and high N_e. In contrast, our empirical data analysis revealed that the relationship between ω_a and π_s weakens when the group average π_s increases, which suggests that long-term fluctuations are not responsible for the positive link between ω_a and π_s we report. Additionally, the gradual decrease of the slope of the relationship with per-group average π_s is also consistent with the fact that the relation is genuine, because (i) we do not expect the regime of demographic fluctuation to correlate with the average π_s of the group, and (ii) there is no relationship between the inter-group variation in π_s and the average π_s of the group (Spearman correlation test: p-value=0.47).

A recently developed method allows the estimation of α and ω_a using polymorphism data alone (20), thus avoiding the assumption of time constancy of the regime of drift and selection. However, estimates of α and ω_a by this method have a meaning that differs from the usual one, as they represent the rate of adaptive evolution of the species during its very recent history, and not the one

of its long-term history. This method requires high quality datasets and highly polymorphic species, and was not applicable to our dataset, in which species and groups differ widely in terms of numbers of SNPs (**Table S3**).

3. Mutation limits adaptation within taxonomic groups in small-θ animals

Our analysis therefore supports the existence of a genuine link between the adaptive substitution rate and θ, which is consistent with the hypothesis that, in several groups of animals, the rate of adaptation is limited by the supply of beneficial mutations. The slope of the relationship is particularly steep in ants, fowls, passerines, rodents and primates (**Figure 2**). For instance, the estimated adaptive rate in rhesus macaque *Macaca mulatta* (π_s=0.0018) is more than three time higher than that of our own species, *Homo sapiens* (π_s=0.0006). Note that this interpretation relies on the assumption that different species from a given taxonomic group share the same DFE and, in particular, the same proportion of beneficial mutations. This result is consistent with previous analyses of the relationship between ω_a and π_s at a relatively recent time scale (27,29). It is also consistent with the finding that strong selective sweeps are more abundant in great apes with large population size (4).

Interestingly, we found that the relationship between ω_a and π_s is significantly stronger in low-diversity than high-diversity groups. In flies, a high-diversity group, the slope of the linear regression between the two variables was only 1.3, whereas it was between 7.8 and 77 in the four groups of vertebrates. In mussels, the taxonomic group with the highest average diversity in our dataset, we detected no significant relationship between ω_a and π_s, the slope being very close to zero (-0.4). It is possible that in such organisms the rate of adaptive evolution is not limited by the supply of beneficial mutations: the standing variation and/or the rate of apparition of new mutations are sufficient for proteins to find the required alleles. This is consistent with the results of (8), who showed that patterns of adaptation to insecticides in natural populations of *Drosophila melanogaster* are incompatible with the hypothesis that adaptation is limited by the income of adaptive mutations. This is also consistent with the results of Jensen and Bachtrog (40), who found very similar rates of adaptation between two *Drosophila* species with different N_e. Finally, this corroborates theoretical predictions indicating that when N_e is sufficiently large, it is the species ability to combine beneficial alleles across loci that limits adaption rather than the strength of

selection or mutation supply (9). Our results suggest that this situation applies to large-N_e groups of animals, such as *Drosophila*, but not to small-N_e ones, such as primates.

4. What are the determinants of ω_a across distantly related taxa?

We used two approaches to estimate the adaptive substitution rate at group level. Both supported a negative among-groups relationship between ω_a and π_s , and between ω_a and life history traits that have been shown to be linked to long-term effective population size (37) (**Figure 1, S2, S3** and **3**). As different sets of genes have been used in the different groups of animals, gene content might influence our results. Indeed, Enard et al. (41) showed that genes interacting with viruses experience a significantly higher adaptive substitution rate, illustrating the importance of gene content. In the exon capture experiment, a subset of genes have been randomly sampled from existing transcriptome reference, whereas in the other species, all available genes (providing that they were present in all species within a group) were used. We do not see any particular reason why the gene sample would be biased with respect to virus interacting proteins in some specific groups, and we did not detect any effect of data type (i.e. exon capture *vs*. genome-wide) on ω_a. Our results are consistent with the results of Galtier (16), who analyzed the relationship between ω_a and π_s in a transcriptomic dataset of 44 distantly related species of animals. Indeed, the main analysis in Galtier (18) revealed no significant correlation between ω_a and π_s, but various control analyses (using GC or expression restricted datasets in particular) yielded a significantly negative correlation between the two variables.

This suggests that the mutation limitation hypothesis does not predict well the variation of ω_a at a large taxonomic scale, i.e., that factors other than θ must be at work here. Below, we discuss a number of such potential factors in the light of Fisher's geometrical model (FGM), which provides a convenient framework under which to think about the determinants of the adaptive substitution rate.

Simulations under FGM and a moving optimum showed that the adaptive substitution rate is primarily determined by the rate of environmental change (32,42). This relates to our result as we may speculate that species with a longer generation time may undergo a higher environmental change per generation, and generation time is (presumably) negatively correlated to population size in animals (37).

Second, Lourenço et al. (32) suggested that organismal complexity, represented by the dimensionality of the phenotypic space in FGM, affects the adaptive substitution rate more strongly than does the effective population size, the adaptive substitution rate being an increasing function of complexity. This is because the probability for a new mutation to be in the optimal direction decreases as the number of potential directions increases, such that the average adaptive walk takes more steps in a high-dimension than a low-dimension space (32,43). Complexity *sensus* FGM, however, is very hard to quantify in a biologically relevant way, and whether primates and birds would be more complex than mussels and worms appears anything but meaningful at organismal level. Different measures of complexity have been thought of at molecular or cellular level, such as genome size, genes or proteins number, number of protein-protein interactions, number of cell types, and these seem to point towards a higher complexity in mammals than insects, for instance (37,38). This would be consistent with the idea of a greater complexity of species with smaller N_e. Fernández and Lynch (46) suggested that the accumulation of mildly deleterious mutations in populations of small size induces secondary selection for protein–protein interactions that stabilize key gene functions, thus introducing a plausible mechanism for the emergence of molecular complexity (46). If the number of protein-protein interactions is a relevant measure of proteome complexity, then this might contribute to explain our report of a higher rate of adaptive substitution in low-θ than in high-θ groups.

Finally, the variation of ω_a across distantly related taxa could be modulated by the long-term N_e via the mean distance of the population to the fitness optimum. Indeed, under FGM, the proportion of beneficial mutations increases with the distance to the optimum. Groups of species evolving under small long-term N_a are further away from their optimum, compared to larger-N_e groups, due to an increased rate of fixation of deleterious mutations at equilibrium, so they are predicted to undergo a larger proportion of beneficial, compensatory mutations. Empirical analyses of SFS based on large samples in humans and flies are consistent with the hypothesis that humans are on average more distant to their optimum than flies (11).

To sum up, our results suggest that factors linked to species long-term effective population size affect the DFE, making the proportion and rate of beneficial mutation non-independent from the long-term N_e. We suggest that the proteome is probably more complex and further away from its optimal state in small-N_e than in large-N_e groups of animals, which might contribute to increase the

steady-state adaptive rate in the former, masking the effect of mutation limitation in across-groups comparisons.

CONCLUSION

In this study, we sampled a large variety of animals species and demonstrated the existence of a timescale-dependent relationship between the adaptive substitution rate and the population mutation rate, that reconciles previous studies that were conducted at different taxonomic scales. We demonstrate that the relationship between the adaptive substitution rate and θ within closely related species sharing a similar DFE is shaped by the limiting supply of beneficial mutations, whereas the between-groups pattern probably reflects the influence of long-term population size on the proportion of beneficial mutations. Our results provide empirical evidence for mutation-limited adaptive rate at whole proteome level in small-N_e groups of animals, and emphasize the fact that the DFE is not independent from the long-term effective population size – a crucial factor that must be properly accounted for in large-scale comparative population genomic analyses.

MATERIAL & METHODS

1. Data set

Genomic, exomic and transcriptomic data from primates, passerines, fowls, rodents and flies were retrieved from the SRA database. Detailed referenced, bioprojects and sample sizes are provided in **Table S1**. The minimal sample size was five diploid individuals (in *Papio anubis*) and the maximum was 20 (in seven species).

Exon capture data were newly generated in ants, butterflies, mussels, earth worms and ribbon worms. We gathered tissue samples or DNA samples for at least eight individuals per species and four or five species per group. Reference transcriptomes were obtained from previously published RNA-seq data in one species per taxonomic group (37,47,48). Details of the species and numbers of individuals are presented in **Table S1**.

2. Multiplexed target capture experiment

DNA from whole animal body (ants), body section (earth worms, ribbon worms), mantle (mussels) or head/thorax (butterflies) was extracted using DNAeasy Blood and Tissue kit (QIAGEN) following the manufacturer instructions. About 3 µg of total genomic DNA were sheared for 20 mn using an ultrasonic cleaning unit (Elmasonic One). Illumina libraries were constructed for all samples following the standard protocol involving blunt-end repair, adapter ligation, and adapter fill-in steps as developed by (49) and adapted in (50).

To perform target capture, we randomly chose contigs in five published reference transcriptomes (*Maniola jurtina* for butterflies (47), *Lineus longissimus* for ribbon worms (37), *Mytilus galloprovincialis* for mussels (37), *Allobophora chlorotica L1* for earth worms (37), and *Formica cunicularia* for ants (48)) in order to reach 2Mb of total sequence length per taxon (~2000 contigs). 100nt-long baits corresponding to these sequences were synthesized by MYbaits (Ann Arbor, MI, USA), with an average cover of 3X.

We then performed multiplexed target capture following the MYbaits targeted enrichment protocol: about 5 ng of each library were PCR-dual-indexed using Taq Phusion (Phusion High-Fidelity DNA Polymerase Thermo Scientific) or KAPA HiFi (2× KAPA HiFi HotStart ReadyMix KAPABIOSYSTEMS) polymerases. We used primers developed in (51). Indexed libraries were purified using AMPure (Agencourt) with a ratio of 1.6, quantified with Nanodrop ND-800, and pooled in equimolar ratio. We had a total of 96 combinations of indexes, and two Illumina lanes, for a total of 244 individuals. This means that we had to index two (rarely three) individuals with the same combination to be sequenced in the same line. When this was necessary, we assignedthe same tag to individuals from distantly related species (i.e. from different groups). Exon capture was achievedaccording to the Mybaits targeted enrichment protocol, adjusting the hybridization temperature to the phylogenetic distance between the processed library and the baits. For libraries corresponding to individuals from the species used to design baits, we used a temperature of 65°C during 22 h. For the other ones we ran the hybridization reactions for 16 h at 65°C, 2 h at 63°C, 2 h at 61°C and 2 h at 59°C. Following hybridization, the reactions were cleaned according to the kit protocol with 200 µL of wash buffers, and hot washes were performed at 65°C or 59°C depending on the samples. The enriched solutions were then PCR-amplified for 14 to 16 cycles, after removal of the streptavidin beads. PCR products were purified using AMPure (Agencourt) with a ratio of

1.6, and paired-end sequenced on two Illumina HiSeq® 2500 lines. Illumina sequencing and demultiplexing were subcontracted.

3. Assembly and genotyping

For RNA-seq data (i.e. fowls and two rodents), we used trimmomatic (52) to remove Illumina adapters and reads with a quality score below 30. We constructed *de novo* transcriptome assemblies for each species following strategy B in (53), using Abyss (54) and Cap3 (55). Open reading frames (ORFs) were predicted using the Trinity package (56). Contigs carrying ORF shorter than 150 bp were discarded. Filtered RNA-seq reads were mapped to this assembly using Burrow Wheeler Aligner (BWA) (version 0.7.12-r1039) (57). Contigs with a coverage across all individual below 2.5xn (where n is the number of individuals) were discarded. Diploid genotypes were called according to the method described in (58) and (59) (model M1) via the software reads2snps (https://kimura.univ-montp2.fr/PopPhyl/index.php?section=tools). This method calculates the posterior probability of each possible genotype in a maximum likelihood framework. Genotypes supported by a posterior probability higher than 95% are retained, otherwise missing data is called. We used version of the method which accounts for between-individual, within-species contamination as introduced in (48), using the -contam=0.1 option, which means assuming that up to 10% of the reads assigned to one specific sample may actually come from a distinct sample, and only validating genotypes robust to this source of uncertainty.

For primates, rodents, passerines and flies, reference genomes, assemblies and annotations files were downloaded from Ensembl (release 89) and NCBI (**Table S1**). We kept only '*CDS*' reports in the annotations files, corresponding to coding exons, which were annotated with the automatic Ensembl annotation pipeline, and the havana team for *Homo sapiens*. We used trimmomatic to remove Illumina adapters, to trim low-quality reads (i.e. with an average base quality below 20), and to keep only reads longer than 50bp. Reads were mapped using BWA (57) on the complete reference assembly. We filtered out hits with mapping quality below 20 and removed duplicates, and we extracted mapping hits corresponding to regions containing coding sequences according to the annotated reference assembly. This was done to avoid calling SNPs on the whole genome, which would be both time consuming and useless in the present context. We called SNPs using a pipeline based on GATK (v3.8-0-ge9d80683). Roughly, this pipeline comprised two rounds of variant calling separated by a base quality score recalibration. Variant calling was first run on every

individuals from every species using HaplotypeCaller (--emitRefConfidence GVCF --genotyping_mode DISCOVERY -hets 0.001). The variant callings from all individuals of a given species were then used to produce a joint genotype using GenotypeGVCFs. Indels in the resulting vcf files were then filtered out using vcftools. The distributions of various parameters associated with SNPs were then used to set several hard thresholds (i.e. Quality by Depth < 3.0; Fisher Strand > 10; Strand Odds Ratio > 3.0; MQRootMeanSquare < 50; MQRankSum < -0.5; ReadPosRankSum < -2.0) in order to detect putative SNP-calling errors using VariantFiltration. This erroneous SNPs were then used for base quality score recalibration of the previously created mapping files using BaseRecalibrator. These mappings with re-calibrated quality scores were then used to re-call variants (HaplotypeCaller), to re-produce a joint genotype (GenotypeGVCFs, --allsites) and to re-set empirical hard thresholds (i.e. same values as above, except for Quality by Depth < 5.0). The obtained vcf files were converted to fasta files (i.e. producing two unphased allelic sequences per individual) using custom python scripts while discarding exons found on both mitochondrial and sexual chromosomes and while filtering out additional SNPs: we removed SNPs with a too high coverage (thresholds were empirically set for each species), with a too low coverage (i.e. 10x per individual) and with a too low genotype quality per individual (i.e. less than 30).

For reads generated through target capture experiment, we cleaned reads with trimmomatic to remove Illumina adapters and reads with a quality score below 30. For each species, we chose the individual with the highest coverage and constructed de novo assemblies using the same strategy as in fowls. Reads of each individuals were then mapped to the newly generated assemblies for each species, using BWA (57). Diploid genotypes were called using the same protocol as in fowls. We used a version of the SNP calling method which accounts for between-individual, within-species contamination as introduced in (48) (see the following section). As the newly generated assemblies likely contained intronic sequences, the predicted cDNAs were compared to the reference transcriptome using blastn searches, with a threshold of e-value of 10e-15. We used an in-house script to remove any incongruent correspondence or inconsistent overlap between sequences from the transcriptomic references and the predicted assemblies, and removed six base pairs at each extremity of the resulting predicted exonic sequences. These high-confidence exonic sequences were used for downstream analyses.

3. Contamination detection and removal

For the newly generated data set, we performed two steps of contamination detection. First, we used the software tool CroCo to detect inter-specific contamination in the *de novo* assembly generated after exon capture (33).

CroCo is a database-independent tool designed to detect and remove cross-contaminations in assembled transcriptomes of distantly related species. This program classifies predicted cDNA in five categories, “clean”, “dubious”, “contamination”, “low coverage” and “high expression”.

Secondly, we used a version of the SNP calling method which accounts for between-individual, within-species contamination as introduced in (48), using the -contam=0.1 option. This means assuming that up to 10% of the reads assigned to one specific sample may actually come from a distinct sample, and only validating genotypes robust to this source of uncertainty.

4. Orthology prediction and divergence analysis

In primates, we extracted one-to-one orthology groups across the six species from the OrthoMaM database (60,61).

In fowls, passerines, rodents and flies, we translated the obtained CDS into proteins and predicted orthology using OrthoFinder (62). In fowls, full coding sequences from the well-annotated chicken genome (Ensembl release 89) were added to the dataset prior to orthology prediction, then discarded. We kept only orthogroups that included all species. We aligned the orthologous sequences with MACSE (Multiple Alignment for Coding SEquences (63).

In each of earth worms, ribbon worms, mussels, butterflies and ants, orthogroups were created via a a blastn similarity search between predicted exonic sequences reference transcriptomes. In each taxon, we concatenated the predicted exonic sequences of each species that matched the same ORF from the reference transcriptome and aligned these using MACSE. We then kept alignments comprising exactly one sequence per species or if only one species was absent.

We estimated lineage specific dN/dS ratio using bppml (version 2.4) and MapNH (version 2.3.2) (64), the former for estimating each branch length and the latter for mapping substitutions on species specific branches.

Tree topologies were obtained from the literature (**Table S4**). In passerines, fowls, rodents, flies and primates, we kept only alignments comprising all the species. In the other groups we also kept alignments comprising all species but one.

We also estimated dN/dS ratios at group level by adding up substitution counts across branches of the trees, including internal branches.

To account for GC-biased gene conversion, we modified the MapNH software such that only GC-conservative substitutions were recorded (26). We estimated the non-synonymous and synonymous number of GC-conservative sites per coding sequence using an in-house script. We could then compute the dN/dS ratio only for GC-conservative substitutions.

5. Polymorphism analysis

For each taxon, we estimated ancestral sequences at each internal node of the tree with the Bio++ program SeqAncestor (64). The ancestral sequences at each internal node were used to orientate single nucleotide polymorphisms (SNPs) of species that descend from this node. We computed non-synonymous (π_n) and synonymous (π_s, i.e. θ) nucleotide diversity, as well as π_n/π_s using the software dNdSpiNpiS_1.0 developed within the PopPhyl project (https://kimura.univ-montp2.fr/PopPhyl/index.php?section=tools) (using gapN_site=4, gapN_seq=0.1 and median transition/transversion ratio values estimated by bppml for each taxonomic group). We also computed unfolded and folded synonymous and non-synonymous site frequency spectra both using all mutations and only GC-conservative mutations using an in-house script as in (18).

6. Mc-Donlad-Kreitman analysis

We estimated α, ω_a and ω_{na} using the approach of (16) as implemented in (18) (program Grapes v.1.0). It models the distribution of the fitness effects (DFE) of deleterious and neutral non-synonymous mutations as a negative Gamma distribution, which is fitted to the synonymous and non-synonymous site frequency spectra (SFS) computed for a set of genes. This estimated DFE is then used to deduce the expected dN/dS under near-neutrality. The difference between observed and expected dN/dS provides an estimate of the proportion of adaptive non-synonymous substitutions, α. The per mutation rate of adaptive and non-adaptive amino-acid substitution were then obtained as following: $\omega_a = \alpha(dN/dS)$ and $\omega_{na} = (1-\alpha)(dN/dS)$. We computed these statistics for each species

using the per branch dN/dS ratio, using either all mutations and substitutions, or only GC-conservative mutations and substitutions.

We used three different distributions to model the fitness effects of mutations that have been shown to perform the best in (18). We then averaged the estimates of the three models using Akaike weights as follows:

$$\overline{\alpha} = \alpha_{GammaZero} * AICw_{GammaZero} + \alpha_{GammaExpo} * AICw_{GammaExpo} + \alpha_{ScaledBeta} * AICw_{ScaledBeta}$$
$$\overline{\omega_a} = \alpha_{GammaZero} * AICw_{GammaZero} + \omega_{aGammaExpo} * AICw_{GammaExpo} + \omega_{aScaledBeta} * AICw_{ScaledBeta}$$
$$\overline{\omega_{na}} = \omega_{naGammaZero} * AICw_{GammaZero} + \omega_{naGammaExpo} * AICw_{GammaExpo} + \omega_{naScaledBeta} * AICw_{ScaledBeta}$$

where AICw stands for akaike weigths that were estimated using the akaike.weights fonction in R (https://www.rdocumentation.org/packages/qpcR/versions/1.4-1/topics/akaike.weights).

When estimating DFE model parameters, we accounted for recent demographic effects by using nuisance parameters, which correct each class of frequency of the synonymous and non-synonymous SFS relative to the neutral expectation in an equilibrium Wright–Fisher population (39).

We also estimated α, ω_a and ω_{na} at group level. Two approaches were used. Firstly, we pooled species specific SFS from each group, and used the dN/dS ratio of the total tree of each taxon. We did so following the unweighted and unbiased strategy of (34), which combines polymorphism data across species with equal weights. Briefly, we divided the synonymous and non-synonymous number of SNPs of each category of the SFS of each species by the total number of SNPs of the species, then we summed those normalized numbers across species and finally we transformed those sums so that the total number of SNPs of the pooled SFS matches the total number of SNPs across species. The resulting estimate was called $\omega_{a[P]}$. Secondly, we calculated the arithmetic mean of ω_{na} across species within a taxonomic group to obtain a non-adaptive substitution rate at the group level. We then subtracted this average from the dN/dS ratio calculating across the whole tree of each taxon to obtain an estimate of the adaptive substitution rate at group level (called $\omega_{a[A]}$).

7. Life history traits variables

Five life history traits were retrieved from the literature for each species: adult size (i.e. the average length of adults), body mass (i.e. the mean body mass of adults' wet weights), fecundity (i.e. the

number of offspring released per day), longevity (i.e. the maximal recorded longevity in years), and propagule size (i.e. the size of the juvenile or egg or larva when leaving parents or group of relatives) (**Table S5**). In the case of social insects and birds, parental care is provided to juveniles until they reach adult size so in these cases, propagule size is equal to adult size.

8. Simulations

In order to evaluate whether our method to estimate the adaptive substitution rate could lead to a spurious correlation between π_s and ω_a, we simulated the evolution of coding sequences in a single population undergoing demographic fluctuations using SLIM V2 (65). We considered panmictic populations of diploid individuals whose genomes consisted of 1500 coding sequences, each of 999 base pairs. We set the mutation rate to 2.2e-9 per base pair per generation, the recombination rate to 10e-8 per base (as in (23)) and the DFE to a gamma distribution of mean -740 and shape 0.14 for the negative part, and to an exponential distribution of mean 10^{-4} for the positive part (those DFE parameters correspond to the DFE estimated from the pooled SFS of primates). We simulated several demographic scenarios with four regimes of frequency of the fluctuations, as well as four regimes of intensity of the fluctuations (see **figure S5**). We sampled polymorphism and divergence for 20 individuals at several time points during the simulations, evaluated π_s and ω_a and measured the correlation between the two variables.

Acknowledgments

We thank Roger de Vila, Lise Dupont, Nicolas Bierne and Christophe Galkowski for providing us with tissue or DNA samples. We thank Iago Bonicci for the homemade program that allowed us to remove intronic sequences of the contigs obtained during the capture experiment by identifying any incongruent correspondence or inconsistent overlap on both the transcriptomic reference and the assembly of the capture experiment contigs. We also thank Yoann Anselmetti, Thibault Leroy and Jonathan Romiguier for their frequent help, and Yoann Anciaux for sharing his experience with Fisher’s geometrical model predictions.

Data accessibility:

Illumina raw reads of the capture experiment are deposited under the Bioproject PRJNA530965 in the SRA database.

Supplementary tables and figures legends:

Table S1 : Details of the species used in this study and numbers of individuals for each species.

Table S2 : Number of orthogroups for each taxonomic group.

The differences in terms of number of orthogroups comes from the fact that we not only kept orthogroups with all species but also orthogroups with all species but one to estimate dN/dS value for each terminal branches in order to maximize the number of substitutions for data sets generated by exon catpure.

Table S3: SNPs counts for each species.

SNPs counts are not integers because they corresponds to SNPs that are present in our SFS, where we chose a sample size (i.e. the number of categories of the SFS) lower that 2*n, where n is the number of individuals. This is to compensate the uneven coverage between individuals that results in some sites in some individuals not to be genotyped. We chose sample sizes that maximize the number of SNPs in each SFS.

Table S4: Sources of the tree topologies of each taxonomic group used to estimate branch length and map substitutions.

Table S5: Values and sources of the life history traits used in this study.

Figure S1: Cross contamination network for de novo assemblies from exon capture.

Circles represent the assemblies, and arrows and their corresponding numbers represent the number of cross contaminants. Most cross contamination events occur between closely-related species and are therefore likely false positive cases.

Figure S2: Relationship between $\omega_{a[P]}$ and π_s and $\log_{10}$ transformed life history traits.

$\omega_{a[P]}$ is estimated using all mutations and substitutions (A) or using only GC-conservative mutations and substitutions (B). Group level π_s and life history traits are estimated by averaging species level estimates across closely related species. The lines represent the regression across groups and * symbols indicate when it is significant.

Figure S3: Relationship between $\omega_{a[A]}$ and π_s and $\log_{10}$ transformed life history traits.

$\omega_a[A]$ is estimated using all mutations and substitutions (A) or using only GC-conservative mutations and substitutions (B). Group level π_s and life history traits are estimated by averaging species level estimates across closely related species. The lines represent the regression across groups and * symbols indicate when it is significant.

Figure S4: Relationship between species-level ω_{na} and π_s and $\log_{10}$ transformed life history traits.

ω_{na} is estimated using all mutations and substitutions (A) or using only GC-conservative mutations and substitutions (B). The dotted lines represent the regression across all species and * symbols indicate when it is significant.

Figure S5: Relationship between species-level α and π_s.

α is estimated using only GC-conservative mutations. The dotted line represents the regression across all species, and full lines represent the regression within each taxonomic groups.

Figure S6: Design of the simulations of fluctuation of population size.

A: three fold ratio between low and high population size and high long-term population size.

B: thirty fold ratio between low and high population size and high long-term population size.

C: three fold ratio between low and high population size and low long-term population size.

D: thirty fold ratio between low and high population size and low long-term population size.

Figure S7: Relationship between ω_a and π_s in simulated scenarios of fluctuating population size.

A: three fold ratio between low and high population size and high long-term population size (scenario A in figure S1)

B: thirty fold ratio between low and high population size and high long-term population size (scenario B in figure S1)

C: three fold ratio between low and high population size and low long-term population size (scenario C in figure S1)

D: thirty fold ratio between low and high population size and low long-term population size (scenario D in figure S1)

References:

1. Bell G. Evolutionary rescue and the limits of adaptation. Philosophical Transactions of the Royal Society B: Biological Sciences. 2013;368(1610):20120080.

2. Lanfear R, Kokko H, Eyre-Walker A. Population size and the rate of evolution. Trends in Ecology & Evolution. janv 2014;29(1):33-41.

3. Smith JM. What Determines the Rate of Evolution? The American Naturalist. 1 mai 1976;110(973):331-8.

4. Nam K, Munch K, Mailund T, Nater A, Greminger MP, Krützen M, et al. Evidence that the rate of strong selective sweeps increases with population size in the great apes. Proceedings of the National Academy of Sciences. 14 févr 2017;114(7):1613-8.

5. Bürger R, Lynch M. Evolution and extinction in a changing environment: a quantitative-genetic analysis. Evolution. 1995;49(1):151-63.

6. Barton N, Partridge L. Limits to natural selection. BioEssays. 2000;22(12):1075-84.

7. Kopp M, Hermisson J. The genetic basis of phenotypic adaptation II: the distribution of adaptive substitutions in the moving optimum model. Genetics. 2009;183(4):1453-76.

8. Karasov T, Messer PW, Petrov DA. Evidence that Adaptation in Drosophila Is Not Limited by Mutation at Single Sites. PLOS Genetics. 17 juin 2010;6(6):e1000924.

9. Weissman DB, Barton NH. Limits to the Rate of Adaptive Substitution in Sexual Populations. PLOS Genetics. 7 juin 2012;8(6):e1002740.

10. Yeaman S, Gerstein AC, Hodgins KA, Whitlock MC. Quantifying how constraints limit the diversity of viable routes to adaptation. PLoS genetics. 2018;14(10):e1007717.

11. Huber CD, Kim BY, Marsden CD, Lohmueller KE. Determining the factors driving selective effects of new nonsynonymous mutations. Proceedings of the National Academy of Sciences. 2017;114(17):4465-70.

12. McDonald JH, Kreitman M. Adaptive protein evolution at the Adh locus in Drosophila. Nature. 20 juin 1991;351(6328):652-4.

13. Smith NGC, Eyre-Walker A. Adaptive protein evolution in *Drosophila*. Nature. févr 2002;415(6875):1022-4.

14. Keightley PD, Eyre-Walker A. Joint inference of the distribution of fitness effects of deleterious mutations and population demography based on nucleotide polymorphism frequencies. Genetics. 2007;177(4):2251-61.

15. Boyko AR, Williamson SH, Indap AR, Degenhardt JD, Hernandez RD, Lohmueller KE, et al. Assessing the evolutionary impact of amino acid mutations in the human genome. PLoS genetics. 2008;4(5):e1000083.

16. Eyre-Walker A, Keightley PD. Estimating the Rate of Adaptive Molecular Evolution in the Presence of Slightly Deleterious Mutations and Population Size Change. Mol Biol Evol. 1 sept 2009;26(9):2097-108.

17. Messer PW, Petrov DA. Frequent adaptation and the McDonald–Kreitman test. PNAS. 21 mai 2013;110(21):8615-20.

18. Galtier N. Adaptive Protein Evolution in Animals and the Effective Population Size Hypothesis. PLOS Genetics. 11 janv 2016;12(1):e1005774.

19. Keightley PD, Campos JL, Booker TR, Charlesworth B. Inferring the Frequency Spectrum of Derived Variants to Quantify Adaptive Molecular Evolution in Protein-Coding Genes of Drosophila melanogaster. Genetics. 1 juin 2016;203(2):975-84.

20. Tataru P, Mollion M, Glémin S, Bataillon T. Inference of Distribution of Fitness Effects and Proportion of Adaptive Substitutions from Polymorphism Data. Genetics. 1 nov 2017;207(3):1103-19.

21. Loewe L, Charlesworth B. Inferring the distribution of mutational effects on fitness in Drosophila. Biology Letters. 22 sept 2006;2(3):426-30.

22. Stoletzki N, Eyre-Walker A. Estimation of the neutrality index. Molecular biology and evolution. 2010;28(1):63-70.

23. Rousselle M, Mollion M, Nabholz B, Bataillon T, Galtier N. Overestimation of the adaptive substitution rate in fluctuating populations. Biology Letters. 1 mai 2018;14(5):20180055.

24. Eyre-Walker A. Changing effective population size and the McDonald-Kreitman test. Genetics. 2002;162(4):2017-24.

25. Corcoran P, Gossmann TI, Barton HJ, Slate J, Zeng K. Determinants of the Efficacy of Natural Selection on Coding and Noncoding Variability in Two Passerine Species. Genome Biol Evol. 1 nov 2017;9(11):2987-3007.

26. Rousselle M, Laverré A, Figuet E, Nabholz B, Galtier N. Influence of Recombination and GC-biased Gene Conversion on the Adaptive and Nonadaptive Substitution Rate in Mammals versus Birds. Mol Biol Evol [Internet]. [cité 22 févr 2019]; Disponible sur: https://academic.oup.com/mbe/advance-article/doi/10.1093/molbev/msy243/5261349

27. Bolívar P, Mugal CF, Rossi M, Nater A, Wang M, Dutoit L, et al. Biased inference of selection due to GC-biased gene conversion and the rate of protein evolution in flycatchers when accounting for it. Mol Biol Evol [Internet]. [cité 3 août 2018]; Disponible sur: https://academic.oup.com/mbe/advance-article/doi/10.1093/molbev/msy149/5063898

28. Ohta T. The Nearly Neutral Theory of Molecular Evolution. Annual Review of Ecology and Systematics. 1992;23(1):263-86.

29. Gossmann TI, Keightley PD, Eyre-Walker A. The Effect of Variation in the Effective Population Size on the Rate of Adaptive Molecular Evolution in Eukaryotes. Genome Biol Evol. 1 janv 2012;4(5):658-67.

30. Gossmann TI, Song B-H, Windsor AJ, Mitchell-Olds T, Dixon CJ, Kapralov MV, et al. Genome wide analyses reveal little evidence for adaptive evolution in many plant species. Molecular biology and evolution. 2010;27(8):1822-32.

31. Strasburg JL, Kane NC, Raduski AR, Bonin A, Michelmore R, Rieseberg LH. Effective population size is positively correlated with levels of adaptive divergence among annual sunflowers. Molecular biology and evolution. 2010;28(5):1569-80.

32. Lourenço JM, Glémin S, Galtier N. The rate of molecular adaptation in a changing environment. Molecular biology and evolution. 2013;30(6):1292-301.

33. Simion P, Belkhir K, François C, Veyssier J, Rink JC, Manuel M, et al. A software tool ‘CroCo’detects pervasive cross-species contamination in next generation sequencing data. BMC biology. 2018;16(1):28.

34. James JE, Piganeau G, Eyre-Walker A. The rate of adaptive evolution in animal mitochondria. Molecular ecology. 2016;25(1):67-78.

35. Galtier N, Roux C, Rousselle M, Romiguier J, Figuet E, Glémin S, et al. Codon Usage Bias in Animals: Disentangling the Effects of Natural Selection, Effective Population Size, and GC-Biased Gene Conversion. Molecular biology and evolution. 2018;35(5):1092-103.

36. Chen J, Glémin S, Lascoux M. Genetic Diversity and the Efficacy of Purifying Selection across Plant and Animal Species. Molecular Biology and Evolution. juin 2017;34(6):1417-28.

37. Romiguier J, Gayral P, Ballenghien M, Bernard A, Cahais V, Chenuil A, et al. Comparative population genomics in animals uncovers the determinants of genetic diversity. Nature. 2014;515(7526):261.

38. Zhen Y, Huber CD, Davies RW, Lohmueller KE. Stronger and higher proportion of beneficial amino acid changing mutations in humans compared to mice and flies. 26 sept 2018 [cité 10 oct 2018]; Disponible sur: http://biorxiv.org/lookup/doi/10.1101/427583

39. Eyre-Walker A, Woolfit M, Phelps T. The Distribution of Fitness Effects of New Deleterious Amino Acid Mutations in Humans. Genetics. 1 juin 2006;173(2):891-900.

40. Jensen JD, Bachtrog D. Characterizing the Influence of Effective Population Size on the Rate of Adaptation: Gillespie's Darwin Domain. Genome Biol Evol. 24 juin 2011;3:687-701.

41. Enard D, Cai L, Gwennap C, Petrov DA. Viruses are a dominant driver of protein adaptation in mammals. :25.

42. Razeto-Barry P, Díaz J, Vásquez RA. The nearly neutral and selection theories of molecular evolution under the fisher geometrical framework: substitution rate, population size, and complexity. Genetics. 2012;191(2):523-34.

43. Orr HA. Adaptation and the Cost of Complexity. Evolution. 1 févr 2000;54(1):13-20.

44. Valentine JW, Collins AG, Meyer CP. Morphological complexity increase in metazoans. Paleobiology. 1994;20(2):131-42.

45. Stumpf MP, Thorne T, de Silva E, Stewart R, An HJ, Lappe M, et al. Estimating the size of the human interactome. Proceedings of the National Academy of Sciences. 2008;105(19):6959-64.

46. Fernández A, Lynch M. Non-adaptive origins of interactome complexity. Nature. 2011;474(7352):502.

47. Rousselle M, Faivre N, Ballenghien M, Galtier N, Nabholz B. Hemizygosity Enhances Purifying Selection: Lack of Fast-Z Evolution in Two Satyrine Butterflies. Genome Biol Evol. 1 oct 2016;8(10):3108-19.

48. Ballenghien M, Faivre N, Galtier N. Patterns of cross-contamination in a multispecies population genomic project: detection, quantification, impact, and solutions. BMC Biology [Internet]. déc 2017 [cité 12 mars 2018];15(1). Disponible sur: http://bmcbiol.biomedcentral.com/articles/10.1186/s12915-017-0366-6

49. Meyer M, Kircher M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harbor Protocols. 2010;2010(6):pdb. prot5448.

50. Tilak M-K, Justy F, Debiais-Thibaud M, Botero-Castro F, Delsuc F, Douzery EJP. A cost-effective straightforward protocol for shotgun Illumina libraries designed to assemble complete mitogenomes from non-model species. Conservation Genet Resour. 1 mars 2015;7(1):37-40.

51. Rohland N, Reich D. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome research. 2012;

52. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20.

53. Cahais V, Gayral P, Tsagkogeorga G, Melo-Ferreira J, Ballenghien M, Weinert L, et al. Reference-free transcriptome assembly in non-model animals from next-generation sequencing data. Molecular ecology resources. 2012;12(5):834-45.

54. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I. ABySS: a parallel assembler for short read sequence data. Genome research. 2009;19(6):1117-23.

55. Huang X, Madan A. CAP3: A DNA sequence assembly program. Genome research. 1999;9(9):868-77.

56. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature biotechnology. 2011;29(7):644.

57. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 2009;25(14):1754-60.

58. Tsagkogeorga G, Cahais V, Galtier N. The population genomics of a fast evolver: high levels of diversity, functional constraint, and molecular adaptation in the tunicate Ciona intestinalis. Genome Biology and Evolution. 2012;4(8):852-61.

59. Gayral P, Melo-Ferreira J, Glémin S, Bierne N, Carneiro M, Nabholz B, et al. Reference-free population genomics from next-generation transcriptome data and the vertebrate–invertebrate gap. PLoS genetics. 2013;9(4):e1003457.

60. Ranwez V, Delsuc F, Ranwez S, Belkhir K, Tilak M-K, Douzery EJ. OrthoMaM: a database of orthologous genomic markers for placental mammal phylogenetics. BMC evolutionary biology. 2007;7(1):241.

61. Douzery EJ, Scornavacca C, Romiguier J, Belkhir K, Galtier N, Delsuc F, et al. OrthoMaM v8: a database of orthologous exons and coding sequences for comparative genomics in mammals. Molecular biology and evolution. 2014;31(7):1923-8.

62. Emms DM, Kelly S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome biology. 2015;16(1):157.

63. Ranwez V, Harispe S, Delsuc F, Douzery EJ. MACSE: Multiple Alignment of Coding SEquences accounting for frameshifts and stop codons. PloS one. 2011;6(9):e22594.

64. Guéguen L, Gaillard S, Boussau B, Gouy M, Groussin M, Rochette NC, et al. Bio++: efficient extensible libraries and tools for computational molecular evolution. Molecular biology and evolution. 2013;30(8):1745-50.

65. Haller BC, Messer PW. SLiM 2: Flexible, Interactive Forward Genetic Simulations. Mol Biol Evol. 1 janv 2017;34(1):230-40.