



Cite this article: Lehtonen J, Lanfear R. 2014
Generation time, life history and the
substitution rate of neutral mutations. *Biol.
Lett.* **10**: 20140801.
<http://dx.doi.org/10.1098/rsbl.2014.0801>

Received: 2 October 2014

Accepted: 6 November 2014

Subject Areas:

evolution, molecular biology

Keywords:

molecular evolution, neutral substitution rate,
life history, generation time, fecundity

Author for correspondence:

Jussi Lehtonen

e-mail: jussi.lehtonen@iki.fi

Electronic supplementary material is available
at <http://dx.doi.org/10.1098/rsbl.2014.0801> or
via <http://rsbl.royalsocietypublishing.org>.

Generation time, life history and the substitution rate of neutral mutations

Jussi Lehtonen^{1,2} and Robert Lanfear^{2,3,4}

¹Evolutionary Biology, Zoological Institute, University of Basel, Vesalgasse 1, 4051 Basel, Switzerland

²Centre of Excellence in Biological Interactions and Division of Evolution Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, Australian Capital Territory 0200, Australia

³National Evolutionary Synthesis Center, Durham, NC 27705-4667, USA

⁴Department of Biological Sciences, Macquarie University, Sydney, New South Wales 2109, Australia

Our understanding of molecular evolution is hampered by a lack of quantitative predictions about how life-history (LH) traits should correlate with substitution rates. Comparative studies have shown that neutral substitution rates vary substantially between species, and evidence shows that much of this diversity is associated with variation in LH traits. However, while these studies often agree, some unexplained and contradictory results have emerged. Explaining these results is difficult without a clear theoretical understanding of the problem. In this study, we derive predictions for the relationships between LH traits and substitution rates in iteroparous species by using demographic theory to relate commonly measured life-history traits to genetic generation time, and by implication to neutral substitution rates. This provides some surprisingly simple explanations for otherwise confusing patterns, such as the association between fecundity and substitution rates. The same framework can be applied to more complex life histories if full life-tables are available.

1. Introduction

Understanding variation in rates and patterns of molecular evolution is an important part of modern biology. In molecular evolution, generation time (T) is expected to be one of the primary determinants of the substitution rate of neutral mutations per time-unit, because species with short T copy their genomes more often, accruing more copy errors per year [1,2]. Although this reasoning is simple, the link between life history (LH) and T is less obvious. For semelparous species, age at first reproduction (b) accurately represents T , because all of the offspring are produced in the first reproductive event. However, for iteroparous species, T cannot be fully described by age at first reproduction, because reproductive events may continue to occur long after the first event.

The most general measure of T , covering semelparity and iteroparity, is the mean age of the parents of a set of new-born individuals (table 1) [3,4]. This defines the average time between reproductive events in a lineage, and the time-frame for the accumulation of errors generated during such events. Theoretical work [5] shows that this definition of T scales the mutation rate per generation (U) to the neutral substitution rate (K) per time unit with a simple equation

$$K = \frac{U}{T}. \quad (1.1)$$

Note that U by itself determines the probability of a new mutation appearing in an offspring [4,6], and therefore lacks a timescale. Because T in iteroparous species is determined by LH-traits such as fecundity and survival, the same must apply to K , independently of variation in U .

Comparative studies of molecular evolution typically compare estimates of LH-traits with K . The set of traits included in these studies varies, as does the

Table 1. Notation, definitions and demographic equations used in the model.

notation	name of parameter, variable or equation	definition
K	neutral substitution rate	rate at which mutations with no effect on fitness are fixed in the genome per time unit
U	mutation probability per generation	the probability that, at a focal site, an offspring has a mutation that its parents do not have
r	population intrinsic growth rate (Malthusian parameter)	the exponential <i>per capita</i> rate of population increase
$f(x)$	fecundity	rate of reproduction in female offspring per mother (age x) per year
b	age at first reproduction	age at which individuals reproduce for the first time
c	length of reproductive time window	time in years during which an individual is reproductively active, if it does not die for stochastic reasons earlier
$d (=b + c)$	age at last reproduction	age in years at which an individual is potentially able to reproduce for the last time, if it has not died for stochastic reasons earlier (close to maximum lifespan in species without menopause)
σ	offspring survival probability	proportion of new-born offspring surviving to reproductive age
μ	adult mortality rate	instantaneous mortality rate of reproductive individuals
$l(x)$	survival probability	probability of survival from birth to age x , taking into account both juvenile and adult survival
σf	recruitment	rate of offspring surviving to maturity per mother per year (i.e. juvenile survival multiplied by fecundity)
T	generation time	mean age of parents, averaged over new-born individuals in a population: $T = \int_0^{\infty} x e^{-rx} l(x) f(x) dx$
	Euler–Lotka equation	$\int_0^{\infty} e^{-rx} l(x) f(x) dx = 1$

definition that is used for T ; for example, age at first reproduction has been interpreted as T in several studies [1,7–10]. However, these simple proxies can fail to account for important sources of variation in T , leading to potentially spurious correlations between K and LH-traits. For example, a correlation between rates of molecular evolution and fecundity was found in mammals [9], even after accounting for variation in other LH-traits. This has been interpreted as evidence for an association between fecundity and the neutral substitution rate, which is independent of generation time, and potential explanations have focused on mechanistic links between fecundity and mutation rates [9,11,12]. However, these hypotheses are not well supported by the available data, and the pattern still lacks a convincing explanation [11]. Furthermore, while one study ruled out a strong association of K with longevity [9], another found the opposite [13], but the latter did not include fecundity as a potential covariate. We suggest that these unexplained results can be partially resolved by properly accounting for LH-effects on T .

2. The model

The equation for T (table 1; [3,4]) requires, in principle, a full description of LH, which is usually not available for all species in comparative studies of molecular evolution. We therefore seek guidelines to predicting and understanding the effect of LH-traits on T and K in iteroparous species, in the absence of detailed data. We use the Euler–Lotka equation

[4,14] (see table 1) to eliminate an unknown LH-parameter from the equation for T . The result is a predicted relationship between the remaining LH-traits and T , and by implication, K .

We assume that individuals survive to age of first reproduction (b) with probability σ , and thereafter have constant rates of mortality (μ) and reproduction (f) until age d . If they reach d , reproduction stops (death may or may not occur at the same time). Age-independent (after first reproduction) μ and f are reasonably realistic for many species [15,16]. In the electronic supplementary material, we show that the main results apply qualitatively even if some of these assumptions are relaxed (a further implicit assumption is that males and females have similar LHs and mutation rates; see [4,5] for methods to account for sex-specificity).

These assumptions imply that for ages $b \leq x \leq d$ survival is $l(x) = \sigma e^{-\mu(x-b)}$, and the reproduction rate is

$$f(x) = \begin{cases} 0 & (\text{for } x < b) \\ f & (\text{for } b \leq x \leq d) \\ 0 & (\text{for } x > d) \end{cases}$$

Substituting into the Euler–Lotka equation yields

$$\begin{aligned} \int_b^d e^{-rx} \sigma e^{-\mu(x-b)} f dx &= \int_b^{b+c} e^{-(r+\mu)x} \sigma e^{\mu b} f dx \\ &= \frac{\sigma f e^{-rb} (1 - e^{-c(r+\mu)})}{r + \mu} = 1, \end{aligned} \quad (2.1)$$

where c is the length of the reproductive time window (i.e. $d = b + c$). Now, recruitment can be solved as a function

of adult mortality, and vice versa

$$\sigma f = \frac{r + \mu}{e^{-rb}(1 - e^{-c(r+\mu)})} \quad (2.2a)$$

and

$$\mu = \frac{y + W(-ye^{-y})}{c} - r, \quad (2.2b)$$

where $y = \sigma f c e^{-rb}$ and W is the principal branch of the Lambert W function [17] (see the electronic supplementary material for details). We then use (2.2a,b) to eliminate either recruitment or mortality from T

$$\begin{aligned} T &= \int_0^\infty x e^{-rx} l(x) f(x) dx = \int_b^{b+c} x e^{-(r+\mu)x} \sigma e^{\mu b} f dx \\ &= b + \frac{\sigma f e^{-rb-c(r+\mu)} (e^{c(r+\mu)} - c(r+\mu) - 1)}{(r+\mu)^2} \\ &= b + \frac{1}{\sigma f e^{-rb}} \left[\frac{y(1 + W(-ye^{-y}))}{y + W(-ye^{-y})} \right], \end{aligned} \quad (2.3a)$$

$$= b + \frac{1}{r + \mu} \left[1 - \frac{ze^{-z}}{1 - e^{-z}} \right], \quad (2.3b)$$

$$= d - \frac{1}{\sigma f e^{-rb}} \left[\frac{y(y-1)}{y + W(-ye^{-y})} \right] \quad (2.3c)$$

and
$$= d - \frac{1}{r + \mu} \left[\frac{z}{1 - e^{-z}} - 1 \right], \quad (2.3d)$$

where $y = \sigma f e^{-rb} c$ and $z = c(r + \mu)$.

Equations (2.3a,b) show how T deviates from its lower limit b , whereas (2.3c,d) show the deviation from the upper limit d . Our interest lies particularly in (2.3a), but the other forms are shown for completeness, and for comparison with previously used approximations of T as a function of adult survival and population growth [18,19].

Examining particular special cases of (2.3a) makes interpretation easier. If we assume that $r \rightarrow 0$ and $c \rightarrow \infty$ (no population growth or effect of senescence on T , reasonable first approximations in long-term evolution under natural conditions), (2.3a) yields

$$T = b + \frac{1}{\sigma f} \quad (2.4)$$

implying

$$K = \frac{U}{T} = \frac{U}{b + (1/\sigma f)} = U \frac{\sigma f}{1 + \sigma f b}. \quad (2.5)$$

Equation (2.5) is an increasing, but saturating function of recruitment σf , compatible with the unexplained empirical observation that K increases with f [9]. The full equation (2.3a) can be seen as (2.4), corrected for maximum lifespan and population growth. Using (2.3a), we can plot estimates of the effect of each LH parameter on the neutral substitution rate (figure 1).

Here, we have used a continuous derivation, as a discrete model does not permit analytical solutions with a limited lifespan (d). However, equation (2.4) is similar if derived with a discrete model.

3. Results and discussion

Equations (2.3) and (2.4) make quantitative predictions for genetic generation time (T) from limited LH data. Because theory [4,5] predicts T to be an important component of

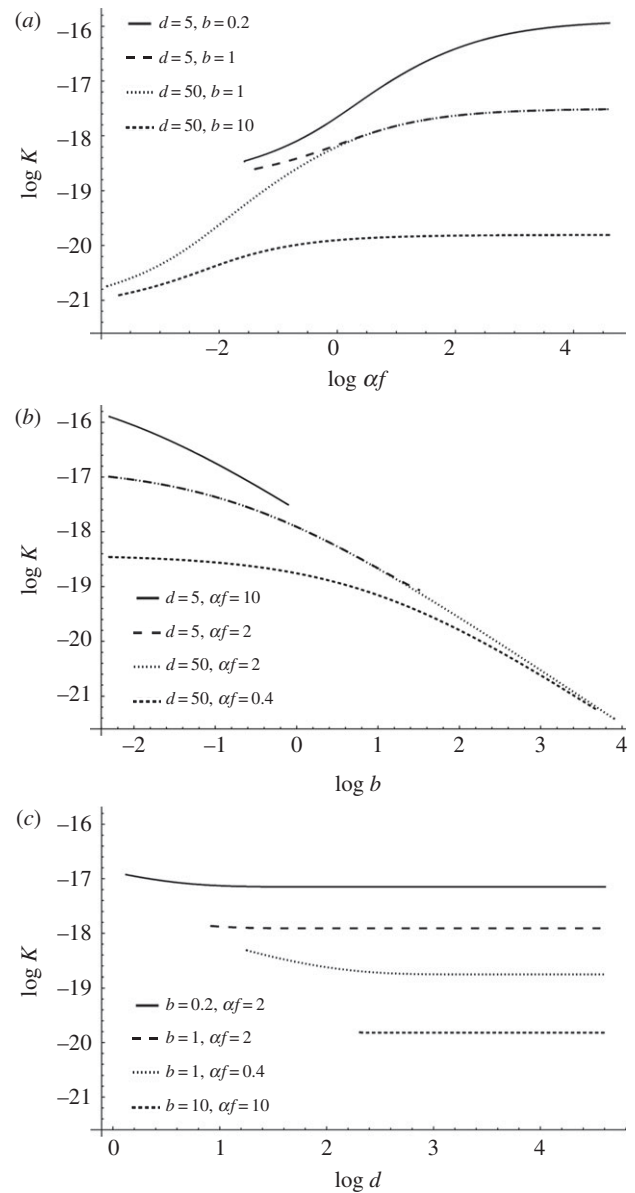


Figure 1. Neutral substitution rate as a function of (a) recruitment, (b) age at first reproduction and (c) age at last reproduction, when the other two variables are held constant. The figure is based on equation (2.3a), with constant population size ($r = 0$), $K = U/T$ and $U = 2.5 \times 10^{-8}$ [20]. The direct causal effect of age at last reproduction on the substitution rate, as predicted by our model, is very small for most parameter combinations (panel (c), but see main text). Not all parameter combinations are possible (e.g. the constraints $b < d$ and $\sigma f(d - b) > 1$ must be fulfilled), which is why some curves are truncated. Displayed combinations illustrate the overall effect of LH-traits on K , but some sections of the curves may be unrealistic, even if possible in principle (e.g. early maturation combined with very low fecundity is unlikely in mammals).

neutral substitution rates (K) in many taxa, these results may both aid future studies of molecular evolution and clarify past studies. For example, our results suggest a simple explanation for the observed correlations between fecundity (f) and K [9]: all else being equal, an iteroparous population with high recruitment (σf) must have high adult mortality, short T (equation (2.4)) and high K (equation (2.5) and figure 1a). More generally, our results show that while recruitment and age at first reproduction (b) strongly influence K (via T), longevity (d) is predicted to have a very minor effect (figure 1c and the electronic supplementary

material). These predictions are broadly in agreement with published comparative studies [1,9,11,21].

A separate question is whether LH traits are correlated with each other. Longevity should evolve in response to traits such as fecundity and age at first reproduction [22] and can also reflect extrinsic mortality [23]. Therefore, longevity alone could still be a reasonable predictor of substitution rates (as found by Lartillot & Delsuc [13]), simply because it is an evolutionary outcome of the same traits that determine T . Combined with the fact that fecundity was not included in [13], this could explain the discrepancies between seemingly similar comparative studies [9,13]. In general, comparative analyses should take into account the possibility that LH traits may evolve in syndromes and are known to often be correlated with each other [11].

The extent to which factors other than T affect K likely varies between taxa and genomes [11,21], potentially making the association between T and K harder to detect empirically. Natural selection may reduce the somatic mitochondrial mutation rate in long-lived species, producing evolutionary correlations between LH and U not accounted

for by our model [24–26]. The link between T and K is also likely to be weakened if germline mutations accumulate throughout the lifespan of the organism, such as occurs in plants [21,27,28] and male mammals [29,30]. Nevertheless, our results have explanatory power to understand substitution rate variation as long as U has a component that does not accumulate continuously throughout an organism's lifetime (e.g. mutations that accumulate during meiosis in each generation; [9]).

Equations (2.1)–(2.5) assume repeated reproduction, and therefore apply only to iteroparous species (e.g. mammals, as studied in references [9,13]). However, equation (1.1) also applies to semelparous species, for which T is simply equal to b , as no further reproduction takes place after the first reproductive event. With semelparity, recruitment (σf) is therefore not expected to affect T or K via the processes described here, which may provide a way to test the theory with data.

Acknowledgements. We thank the referees for helpful comments.

Funding statement. J.L. was supported by the Kone Foundation and R.L. by the Australian Research Council.

References

1. Thomas JA, Welch JJ, Lanfear R, Bromham L. 2010 A generation time effect on the rate of molecular evolution in invertebrates. *Mol. Biol. Evol.* **27**, 1173–1180. (doi:10.1093/molbev/msq009)
2. Li W-H, Ellsworth DL, Krushkal J, Chang BHJ, Hewett-Emmett D. 1996 Rates of nucleotide substitution in primates and rodents and the generation–time effect hypothesis. *Mol. Phylogenet. Evol.* **5**, 182–187. (doi:10.1006/mpev.1996.0012)
3. Leslie P. 1966 The intrinsic rate of increase and the overlap of successive generations in a population of Guillemots (*Uria aalge* Pont.). *J. Anim. Ecol.* **35**, 291–301. (doi:10.2307/2396)
4. Charlesworth B. 1994 *Evolution in age-structured populations*. Cambridge, UK: Cambridge University Press.
5. Pollak E. 1982 The rate of mutant substitution in populations with overlapping generations. *Genet. Res.* **40**, 89–94. (doi:10.1017/S0016672300018930)
6. Gillespie JH. 2010 *Population genetics: a concise guide*. Baltimore, MD: JHU Press.
7. Mooers AØ, Harvey PH. 1994 Metabolic rate, generation time, and the rate of molecular evolution in birds. *Mol. Phylogenet. Evol.* **3**, 344–350. (doi:10.1006/mpev.1994.1040)
8. Gissi C, Reyes A, Pesole G, Saccone C. 2000 Lineage-specific evolutionary rate in mammalian mtDNA. *Mol. Biol. Evol.* **17**, 1022–1031. (doi:10.1093/oxfordjournals.molbev.a026383)
9. Welch J, Bininda-Emonds O, Bromham L. 2008 Correlates of substitution rate variation in mammalian protein-coding sequences. *BMC Evol. Biol.* **8**, 53. (doi:10.1186/1471-2148-8-53)
10. Tsantes C, Steiper ME. 2009 Age at first reproduction explains rate variation in the strepsirrhine molecular clock. *Proc. Natl Acad. Sci. USA* **106**, 18 165–18 170. (doi:10.1073/pnas.0906686106).
11. Bromham L. 2011 The genome as a life-history character: why rate of molecular evolution varies between mammal species. *Phil. Trans. R. Soc. B* **366**, 2503–2513. (doi:10.1098/rstb.2011.0014)
12. Britten RJ. 1986 Rates of DNA-sequence evolution differ between taxonomic groups. *Science* **231**, 1393–1398. (doi:10.1126/science.3082006)
13. Lartillot N, Delsuc F. 2012 Joint reconstruction of divergence times and life-history evolution in placental mammals using a phylogenetic covariance model. *Evolution* **66**, 1773–1787. (doi:10.1111/j.1558-5646.2011.01558.x)
14. Lotka AJ. 1925 *Elements of physical biology*. Baltimore, MD: Williams & Wilkins.
15. Lande R, Engen S, Saether B-E. 2003 *Stochastic population dynamics in ecology and conservation*. Oxford, UK: Oxford University Press.
16. Charlesworth B. 2001 The effect of life-history and mode of inheritance on neutral genetic variability. *Genet. Res.* **77**, 153–166. (doi:10.1017/S0016672301004979)
17. Corless RM, Gonnet GH, Hare DE, Jeffrey DJ, Knuth DE. 1996 On the Lambert W function. *Adv. Comput. Math.* **5**, 329–359. (doi:10.1007/BF02124750)
18. Lande R. 1988 Demographic models of the northern spotted owl (*Strix occidentalis caurina*). *Oecologia* **75**, 601–607. (doi:10.1007/BF00776426)
19. Gaillard JM, Yoccoz NG, Lebreton JD, Bonenfant C, Devillard S, Loison A, Pontier D, Allaine D. 2005 Generation time: a reliable metric to measure life-history variation among mammalian populations. *Am. Nat.* **166**, 119–123. (doi:10.1086/430330)
20. Nachman MW, Crowell SL. 2000 Estimate of the mutation rate per nucleotide in humans. *Genetics* **156**, 297–304.
21. Gaut B, Yang L, Takuno S, Eguiarte LE. 2011 The patterns and causes of variation in plant nucleotide substitution rates. *Annu. Rev. Ecol. Syst.* **42**, 245–266. (doi:10.1146/annurev-ecolsys-102710-145119)
22. Hamilton WD. 1966 Moulding of senescence by natural selection. *J. Theor. Biol.* **12**, 12–45. (doi:10.1016/0022-5193(66)90184-6)
23. Abrams PA. 1993 Does increased mortality favor the evolution of more rapid senescence? *Evolution* **47**, 877–887. (doi:10.2307/2410191)
24. Galtier N, Jobson RW, Nabholz B, Glémin S, Blier PU. 2009 Mitochondrial whims: metabolic rate, longevity and the rate of molecular evolution. *Biol. Lett.* **5**, 413–416. (doi:10.1098/rsbl.2008.0662)
25. Nabholz B, Glémin S, Galtier N. 2008 Strong variations of mitochondrial mutation rate across mammals—the longevity hypothesis. *Mol. Biol. Evol.* **25**, 120–130. (doi:10.1093/molbev/msm248)
26. Nabholz B, Glémin S, Galtier N. 2009 The erratic mitochondrial clock: variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. *BMC Evol. Biol.* **9**, 54. (doi:10.1186/1471-2148-9-54)
27. Lanfear R, Ho SY, Davies TJ, Moles AT, Aarssen L, Swenson NG, Warman L, Zanne AE, Allen AP. 2013 Taller plants have lower rates of molecular evolution. *Nat. Commun.* **4**, 1879. (doi:10.1038/ncomms2836)
28. Klekowski EJ, Godfrey PJ. 1989 Ageing and mutation in plants. *Nature* **340**, 389–391. (doi:10.1038/340389a0)
29. Kong A *et al.* 2012 Rate of *de novo* mutations and the importance of father's age to disease risk. *Nature* **488**, 471–475. (doi:10.1038/nature11396)
30. Thomas GWC, Hahn MW. 2013 The human mutation rate is increasing, even as it slows. *Mol. Biol. Evol.* **31**, 253–257. (doi:10.1093/molbev/mst118)