Title

Response to Bakker et al.

Permalink

https://escholarship.org/uc/item/98w8g7w4

Journal

Current Biology, 32(8)

ISSN

0960-9822

Authors

Robinson, Jacqueline A
Bowie, Rauri CK
Dudchenko, Olga
et al.

Publication Date

2022-04-01

DOI

10.1016/j.cub.2022.03.049

Peer reviewed

Effect of varying generation time and mutation rate parameters on demographic inference: Response to Bakker et al.

Jacqueline A. Robinson[1*†], Rauri C. K. Bowie[2,3], Olga Dudchenko[4,5], Erez Lieberman Aiden[4,5,6,7] Sher L. Hendrickson[8,9], Cynthia C. Steiner[10], Oliver A. Ryder[10,11], David P. Mindell[3], Jeffrey D. Wall[1*]

[1]Institute for Human Genetics, University of California, San Francisco, CA, USA
[2]Department of Integrative Biology, University of California, Berkeley, CA, USA
[3]Museum of Vertebrate Zoology, University of California, Berkeley, CA, USA
[4]The Center for Genome Architecture, Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA
[5]Center for Theoretical and Biological Physics, Rice University, Houston, TX, USA
[6]Shanghai Institute for Advanced Immunochemical Studies, ShanghaiTech, Pudong, China
[7]Faculty of Science, UWA School of Agriculture and Environment, University of Western Australia, Perth WA, Australia
[8]Department of Biology, Shepherd University, Shepherdstown, WV, USA
[9]Center for Conservation Genomics, Smithsonian Conservation Biology Institute, Washington, DC, USA
[10]San Diego Zoo Wildlife Alliance, Beckman Center for Conservation Research, Escondido, CA, USA
[11]Department of Evolution, Behavior, and Ecology, University of California, San Diego, CA, USA
*Correspondence: jacqueline.robinson@ucsf.edu
†Lead contact

The diversity within a diploid genome (i.e. heterozygosity) reflects long-term effective population size (N_e) over evolutionary timescales, and can therefore be used to estimate species historical abundance (e.g. [1,2]). Doing so requires knowledge about mutation rate (μ) and generation time (G), parameters that are challenging to estimate, that may be correlated with one another, and are often unknown for a given species. Bakker et al.[3] contend that our[4] use of the same G to infer ancestral N_e of California condors (*Gymnogyps californianus*), Andean condors (*Vultur gryphus*), and turkey vultures (*Cathartes aura*) produces unreliable results given these species' different contemporary generation times. They further argue that inaccurate N_e estimates from these methods could adversely affect the conservation of threatened species. We agree with Bakker et al. that assumptions about G are crucial to inferring demographic history from genetic data and should be carefully considered. However, we disagree with their assertions regarding estimates of N_e that neglect the interrelatedness of G and μ, and the resulting implications. As we demonstrate below, if these closely related species share similar rates of evolution (germline mutations per generation), then the California condor was previously more abundant than the turkey vulture, regardless of G.

40

41   Our study revealed that California condors have the highest genome-wide diversity among all

42   three species after excluding runs of homozygosity caused by recent inbreeding, which suggests

43   they were once more abundant than Andean condors and turkey vultures[4]. Results from

44   demographic inference with PSMC/MSMC[1,2] support this conclusion when the same μ and G

45   are used for all three species. The G parameter rescales time from generations into years, and

46   solely affects the timescale of the inference (x-axis), whereas μ affects both the timescale and

47   N_e (y-axis). As noted previously[3,4], neither parameter affects the shape of the inferred

48   demographic trajectory. In reality, both μ and G can evolve and change over time, but only a

49   constant value may be provided for rescaling PSMC/MSMC output. As the rate of evolution in

50   condors and turkey vultures is unknown, our study provided results under multiple

51   combinations of G and μ, and gave the rationale for our parameter choices, which were based

52   on values from prior studies and genome-wide divergence in our dataset[4]. However, in the

53   absence of information on how evolutionary rate differs between the species in our study, we

54   did not explore the effects of varying G and μ between condors and turkey vultures. Below, we

55   now demonstrate how violating the assumption of shared μ and G affects our inference of

56   ancestral N_e, and that the central issue is whether these closely related species share μ per

57   *generation* or per *year*.

58

59   Per-generation mutation rates ranging from 0.20 to 1.66×10^{-8}/site/generation have recently

60   been estimated using counts of *de novo* mutations in ~20 vertebrate species[5], but only one of

61   these is from an avian species: Smeds et al.[6] estimated a rate of 4.6×10^{-9}/site/generation in the

62   collared flycatcher (*Ficedula albicollis*), a passerine. A different approach uses time-calibrated

63   phylogenies, but yields estimates of mutation rate per year, rather than per generation. Using

64   this method, Zhang et al.[7] found that avian substitution rates vary from 1.0 to 3.3×10^{-9}

65   substitutions/site/year, with the lowest rate in Struthioniformes (ostriches) and the highest in

66   Passeriformes (passerines), consistent with the longstanding observation that mutation rates

67   typically correlate with life history traits such as adult body size and generation time (e.g. [8,9]). In

68   other words, G and μ are interrelated, and changes in G may conceivably effect changes in μ, or

69   vice versa. California and Andean condors represent extremes among birds in terms of body

70   size and longevity, making it difficult to extrapolate μ from earlier studies. However, the turkey

71   vulture is less of an outlier. Zhang et al.[7] estimated a substitution rate in Accipitriformes (the

72   order including New World vultures) of ~1.4×10^{-9}/site/year. On average, G in this group is ~7.5

73   years[10], close to the value of 7 years preferred by Bakker et al.[3] for turkey vultures, yielding a

74   rate of ~1×10^{-8}/site/generation, squarely within the range estimated in other vertebrates[5].

75   However, multiplying 1.4×10^{-9}/site/year by 25 years, a value of G for condors preferred by

76   Bakker et al.[3], yields a mutation rate of 3.5×10^{-8}/site/generation, far exceeding previous

77   estimates. Potentially, long generation times and large body size in condors may be coupled

78   with a μ per year below what has been estimated previously[7], yielding a value of μ per

79   generation more in line with other species.

80

81   Here, we demonstrate the effects of varying μ and G on our demographic inference, using

82   values of G preferred by Bakker et al.[3]. Fig. 1 A and B reproduce results from our study for G=10

83   years and two values of μ (0.65×10^{-9} and 1.40×10^{-9}/site/year; see Robinson et al.[4] for details).

84   Here, all three species share the same parameter values, and the California condor has the

85   highest ancestral N_e among all three species. Changing G while keeping μ per generation the

86   same shifts trajectories along the x-axis, as shown in Fig. 1 C and D, but has no effect on

87   estimates of N_e. In this scenario, California condors were still more abundant than turkey

88   vultures at one time, but the turkey vulture becomes the most abundant species over the past

89   100K years. In a third scenario where instead, all species share the same μ per year, changing G

90   also affects μ per generation, as shown in Fig. 1 E and F. Here, the turkey vulture is consistently

91   more abundant than either condor species. Thus, if these closely related species share a similar

92   μ per generation, California condors were historically more abundant than turkey vultures

93   >>10K years ago, regardless of G, but not if μ per generation in turkey vultures is substantially

94   lower. Future studies of avian mutation rates may shed light on which scenario is more likely.

95

96   Our PSMC/MSMC results are accessible (https://doi.org/10.5281/zenodo.4680008) so that

97   figures can be re-plotted with different parameters, as demonstrated above. Finally, we advise

98   against the use of PSMC/MSMC results for setting conservation targets, partly because of the

99   limitations of these methods, but also because ancient population sizes from >10K years ago

100   are rarely relevant in the context of contemporary conservation. As we cautioned previously,

101   PSMC/MSMC provides no information about N_e over recent timescales. Regardless of the

102   values of μ and G in condors and turkey vultures, historical N_e in these species was clearly

103   dynamic and very different from what it is today, as shown here and in our original study.

104

Author contributions

106   J.A.R. performed analyses and wrote the manuscript. All authors edited and approved the

107   manuscript.

108

Declaration of Interests

110   J.D.W. receives research funding from Sierra Pacific Industries. E.L.A. is Scientific Advisory Board

111   co-chair and consultant for HolyHaid Lab Corporation (Shenzhen, China), whose parent

112   company is Hollyhigh International Capital (Beijing & Shanghai, China).

113

Figure Legends

115   **Fig. 1. Effect of varying G and μ parameters on demographic inference with PSMC.** Generation

116   times (in years) shown in parentheses next to each sample. Note the units of μ (per generation

117   or per year). A, B) Results reproduced from Robinson et al.[4] under μ=0.65x10^{-9} and

118   1.40x10^{-9}/site/year and G=10 years for all three species. C, D) Changing G while keeping the

119   same per-generation μ as in (A, B) shifts trajectories along the x-axis, but has no effect on

120   estimates of N_e. E, F) Alternatively, changing G while keeping the same per-year μ as in (A, B)

121   causes the per-generation μ to vary between species, resulting in changes to N_e (y-axis).

122

References

1. Li, H. and Durbin, R., 2011. Inference of human population history from individual whole-genome sequences. *Nature*, 475(7357), 493-496.

2. Schiffels, S. and Durbin, R., 2014. Inferring human population size and separation history from multiple genome sequences. *Nature Genetics*, 46(8), 919-925.

3. Bakker, V.J., Finkelstein, M.E., D'Elia, J.D., Doak, D.F., Kirkland, S., 2021. Genetically based demographic reconstructions require careful consideration of generation time. *Current Biology*.

4. Robinson, J.A., Bowie, R.C., Dudchenko, O., Aiden, E.L., Hendrickson, S.L., Steiner, C.C., Ryder, O.A., Mindell, D.P. and Wall, J.D., 2021. Genome-wide diversity in the California condor tracks its prehistoric abundance and decline. *Current Biology*, 31(13), 2939-2946.

5. Bergeron, L.A., Besenbacher, S., Turner, T., Versoza, C.J., Wang, R.J., Price, A.L., Armstrong, E., Riera, M., Carlson, J., Chen, H.Y. Hahn, M.W. et al. 2022. The mutationathon highlights the importance of reaching standardization in estimates of pedigree-based germline mutation rates. *Elife*, 11, p.e73577.

6. Smeds, L., Qvarnström, A. and Ellegren, H., 2016. Direct estimate of the rate of germline mutation in a bird. *Genome Research*, 26(9), 1211-1218.

7. Zhang, G., Li, C., Li, Q., Li, B., Larkin, D.M., Lee, C., Storz, J.F., Antunes, A., Greenwold, M.J., Meredith, R.W., Ödeen, A. et al., 2014. Comparative genomics reveals insights into avian genome evolution and adaptation. *Science*, 346(6215), 1311-1320.

8. Mooers, A.Ø. and Harvey, P.H., 1994. Metabolic rate, generation time, and the rate of molecular evolution in birds. *Molecular Phylogenetics and Evolution*, 3(4), 344-350.

9. Welch, J.J., Bininda-Emonds, O.R. and Bromham, L., 2008. Correlates of substitution rate variation in mammalian protein-coding sequences. *BMC Evolutionary Biology*, 8(1), 1-12.

10. Bird, J.P., Martin, R., Akçakaya, H.R., Gilroy, J., Burfield, I.J., Garnett, S., Symes, A., Taylor, J., Şekercioğlu, Ç.H. and Butchart, S.H., 2020. Generation lengths of the world's birds and their implications for extinction risk. *Conservation Biology*, 34(5), 1252-1261.