Background selection theory overestimates effective population size for high mutation rates

Joseph Matheson[1], Joanna Masel[1]

[1]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, 85721, USA

Corresponding author: Joanna Masel

ABSTRACT

Simple models from the neutral theory of molecular evolution are claimed to be flexible enough to incorporate the complex effects of background selection against linked deleterious mutations. Complexities are collapsed into an "effective" population size that specifies neutral genetic diversity. To achieve this, current background selection theory assumes linkage equilibrium among deleterious variants. Data do not support this assumption, nor do theoretical considerations when the genome-wide deleterious mutation is realistically high. We simulate genomes evolving under background selection, allowing the emergence of linkage disequilibria. With realistically high deleterious mutation rates, neutral diversity is much lower than predicted from previous analytical theory.

Keywords: linkage disequilibrium, population genetics, nearly neutral theory, forward time simulation, effective population size, expected heterozygosity

INTRODUCTION

The neutral theory of molecular evolution postulates that i) most genetic diversity observed in natural populations is neutral with respect to an organism's fitness (Kimura 1968; King and Jukes 1969), and ii) dynamics are well-described by models of a single neutral locus in a population of a specified "effective" population size (Ewens 2004; Charlesworth 2009; Masel 2011; Kern and Hahn 2018). This elegant mathematical framework has since been expanded to incorporate migration among populations (Wang

and Whitlock 2003), temporal changes in the effective population size (Wright 1938; Vucetich *et al.* 2017), a threshold for neutrality that varies among populations (Ohta 1973), and the effects of selection against deleterious mutations on linked neutral variants (Charlesworth *et al.* 1993). Approximating complex genomic dynamics with single-locus neutral models is extraordinarily powerful, but how accurate is this approximation? Here we focus on the case of background selection at realistically many sites.

The original version of neutral theory modeled one neutral locus in a population of constant size, idealized to obey either Wright-Fisher or Moran dynamics. With more complex demographic histories, an effective population size N_e can be defined as the size of an idealized population that produces the same value of a chosen population statistic (Charlesworth 2009). The statistic usually chosen is genetic diversity (expected heterozygosity) to produce the coalescent effective population size. In an idealized population, genetic diversity depends only on the mutation rate at that locus and the census size of the population (Kimura 1969).

Not long after neutral theory was proposed, it became clear that it was incompatible with some patterns of molecular evolution, in particular the independence of rates of evolution from generation time (Ohta 1973). Neutral theory was therefore replaced by nearly neutral theory. Neutral theory considers mutations that either are strictly neutral or so deleterious that they can be ignored. Nearly neutral theory retains the binary distinction between rapidly purged versus neutral mutations, but allows the ratio of mutations across these two types to vary among species. This ratio is determined by another effective population size, sometimes referred to today as the drift barrier effective population size (Ohta 1973; Sung *et al.* 2012). Nearly neutral theory still derives genome-wide patterns of diversity from models of many independent single loci.

The problem with this binary distinction is that slightly deleterious mutations are purged only slowly from populations. During this removal process, they depress genetic variation at linked sites in the genome, a phenomenon known as background selection (Charlesworth *et al.* 1993). The depression in genetic

variation caused by background selection is typically modeled as a decrease in the coalescent effective population size for the neutral loci linked to deleterious variants (Hudson and Kaplan 1995; Lohmueller *et al.* 2011; Comeron 2014). In a population with no recombination, the coalescent effective population size would decrease from N_e to $f_0 N_e$, where f_0 is the equilibrium frequency of individuals with no deleterious mutations, because any neutral variants linked to deleterious variants would be doomed (Charlesworth *et al.* 1993).

Recombination can decouple neutral variants from deleterious variants, resulting in less depression of variation under background selection (Cutter and Payseur 2013). For a single neutral locus linked to a single locus where deleterious mutations with fixed effect size s occur at rate u per diploid individual per generation, and with recombination between the loci occurring at rate r, heterozygosity at the neutral locus is reduced by a factor $F \approx 1 - \frac{us}{2(s+r)^2}$ (see Figure 1A) (Hudson and Kaplan 1994). This result can be straightforwardly extended to any number of deleterious sites linked to the focal neutral site by assuming that mutation and recombination rates are uniform across a genomic window, and that there is linkage equilibrium among deleterious variants e.g. because multiple significantly linked deleterious mutations are not present at the same time (see Figure 1B) (Hudson and Kaplan 1995; Nordborg *et al.* 1996). In this case, the ratio of observed N_e (based on heterozygosity) to N at a focal neutral site in a genomic window of any size is given by $\frac{N_e}{N} = e^{-U_w/2s+R_w}$, where U_w is now the total diploid deleterious mutation rate across the entire window, and R_w is now the total recombination rate between the ends of the window (Hudson and Kaplan 1995). Since $R_w \gg s$, this can be approximated as $\frac{N_e}{N} = e^{-U_w/R_w}$, producing the result that the effect of background selection on N_e can be approximated with a single factor that depends only on rates of deleterious mutation and recombination. With the same assumptions, similar results can be obtained in the case where both deleterious and beneficial mutations occur (Kim and Stephan 2000).

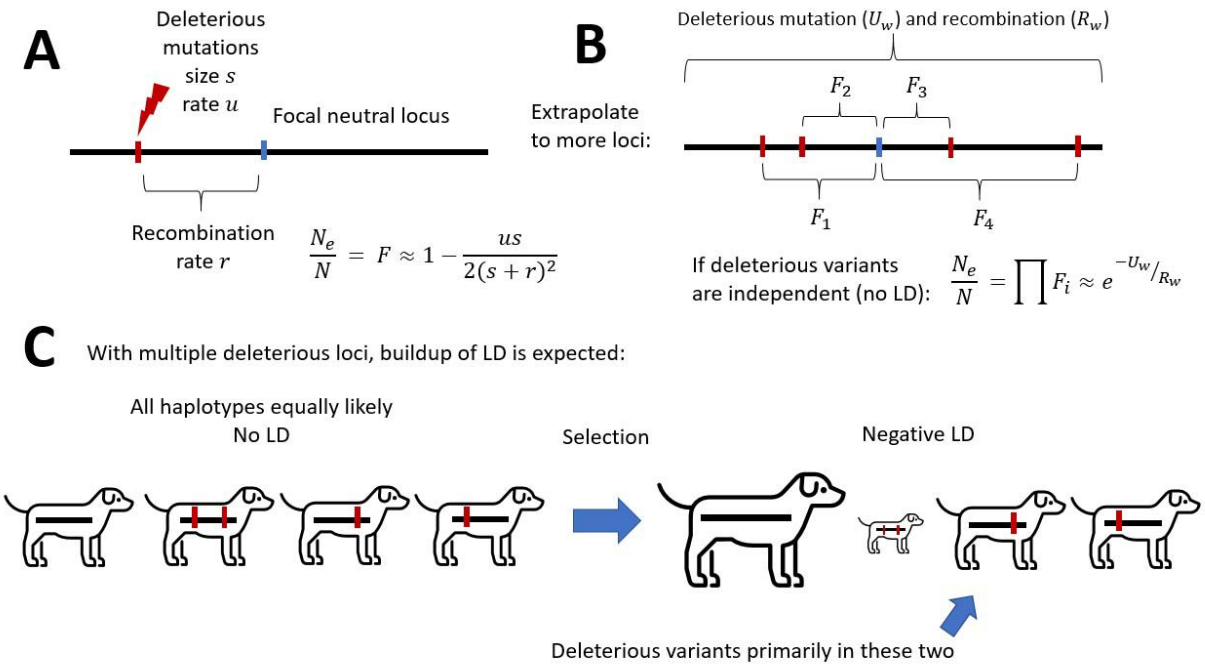

Figure 1. Analytical approximations of the effects of background selection on neutral diversity depend on an unjustified assumption of linkage equilibrium among deleterious variants. (A) The reduction in variation at a focal neutral locus linked to a single deleterious locus can be solved as a function of the deleterious mutation rate, selective effect of deleterious mutations, and recombination rate between the two loci. (B) This result can be extended to a focal neutral locus in a genomic window with a deleterious mutation rate and recombination rate specified across the whole window rather than at a single site. This requires the assumption of linkage equilibrium among deleterious variants. (C) However, linkage disequilibrium is expected between deleterious variants. Left: for two deleterious loci, we initialize all four haplotypes at linkage equilibrium. Selection will then tend to destroy positive linkage disequilibria by removing haplotypes with both deleterious variants and promoting haplotypes with no deleterious variants. Right: after selection, deleterious variants will therefore be over-represented by haplotypes with one or the other variant but not both, producing net negative linkage disequilibrium.

But an excess of negative linkage disequilibrium is expected among selected mutations (Barton and Otto 2005; Keightley and Otto 2006), violating a key assumption of the Hudson and Kaplan (Hudson and Kaplan 1995) model. Each new mutation is born into either positive or negative linkage disequilibrium with each previously circulating mutation, and these disequilibria tend to be amplified by subsequent

random genetic drift when physical linkage keeps recombination low. The two kinds of disequilibria will cancel out across a statistical average when only mutation and drift are considered, but selection quickly eliminates positive disequilibria, either by removing haplotypes with multiple deleterious variants or by fixing haplotypes with multiple beneficial variants (see Figure 1C), leaving an excess of negative disequilibria (Hill and Robertson 1968; Felsenstein 1974). Linkage disequilibrium has been found to enhance the loss of neutral variation in the somewhat different case of selective sweeps (Barton 1998). This suggests that Hudson and Kaplan's approximation might similarly underestimate the effect of background selection on neutral diversity.

Violation of the assumption of linkage equilibrium among deleterious variants might not be a trivial matter given the sheer quantity of deleterious mutations entering populations. For example, the average human is estimated to begin life with an average of two new deleterious mutations not present in either parent (Lesecque *et al.* 2012), and high deleterious mutation rate estimates are not unique to humans (Haag-Liautard *et al.* 2007). Some argue that deleterious mutation rates are even higher, closer to 10 new deleterious mutations per person in humans (Kondrashov 2017).

Here we perform a multi-locus simulation using the fwdpy package (Thornton 2014) that efficiently handles large numbers of non-neutral mutations in relatively large census size populations (Haller and Messer 2017). The recent addition of tree-sequence recording (Kelleher *et al.* 2018) to fwdpy additionally allows the calculation of a coalescent effective population size without needing to explicitly simulate and track neutral mutations. This approach allows us to compare the 'observed' coalescent effective population size from our high-deleterious-mutation-rate simulations to the analytic expectations from (Hudson and Kaplan 1995)'s model that assumes linkage equilibrium among deleterious mutations. More broadly, this can inform whether the general approach of single locus models of a neutral mutation are appropriate for a population subject to background selection under realistically high deleterious mutation rates and the resulting linkage disequilibrium among deleterious mutations.

METHODS

Multi-locus simulations

All simulations were written in Python using fwdpy (Thornton 2014). We simulated populations of N diploid individuals undergoing selection against deleterious mutations using a standard Wright-Fisher model for $10N$ generations. Each individual's genome was made up of 23 chromosomes of length 100 under an infinite-sites model (i.e., all floating-point numbers between 0 and 100 on each chromosome are potential loci). Recombination occurs by crossing-over exactly twice per chromosome, matching data for humans (Pardo-Manuel De Villena and Sapienza 2001), although we allow recombination to occur anywhere rather than explicitly simulating a centrosome or recombination hotspots.

Deleterious mutations occur with genome-wide rate U, and have fixed selection coefficients s. In the "no genes" condition, they are located uniformly at random along the chromosomes, while in the "genes" condition they occur only in "genes". We simulate 1,000 genes, accounting for 10 percent of the genome, interspersed at regular intervals throughout the genome. These parameters were not chosen to be representative of any particular species, but simply to capture the qualitative consequences of clustering among sites subject to background selection.

A recent study of a large sample of modern European humans estimated a gamma distribution of fitness effects of new mutations with mean $sN_e = -224.33$ and $N_e = 23{,}646$, implying a mean $s \approx$ -0.01 (Kim *et al.* 2017). In our main results, we simplify to use a constant s = -0.01 to avoid complications from slightly deleterious mutations with sN_e near 1. We also explore higher and lower values of s, and the complete distribution with the reference mean.

While our forward-time simulations track only deleterious mutations, we compute genetic diversity and hence effective population size by using tree-sequence recording (Kelleher *et al.* 2018) during the simulation, which allows neutral mutations to be projected backwards onto the genealogical histories of

different genomic regions. Neutral mutations occur uniformly at random on the entire genome at an arbitrary rate 10^{-4} per genomic 'unit', for a total rate of 0.23 per genome. This low value provides sufficient resolution of N_e at low computational cost. We use msprime (Kelleher *et al.* 2016) to calculate neutral diversity θ on the resulting tree sequence, and then calculate the effective population size for a simulation using $\theta = 4N_e\mu$ and solving for N_e.

We simulated census population sizes *N* ranging from 2000 to 10,000. This is compatible with the range of inferred estimates for human effective population sizes (Takahata 1993; Tenesa *et al.* 2007; McEvoy *et al.* 2011). In all cases, we compared the values of $\frac{N_e}{N}$ calculated from the neutral diversity in the multi-locus simulations to the expected $\frac{N_e}{N}$ ratio given from the two-locus model.

RESULTS

Multi-locus simulations produce a lower N_e than expected from two-locus analytical predictions (Figure 2). The discrepancy becomes marked when the genome-wide mutation rate U is high, specifically when $U > 1$, as is estimated to be the case for humans (Lesecque *et al.* 2012).

This main result is independent of *N* and *s*. A five-fold change in census population size *N* has no significant effect on the $\frac{N_e}{N}$ ratio (Figure 2A). Larger selective effect sizes only slightly increase the degree to which $\frac{N_e}{N}$ drops with *U* (Figure 2B). Simulating a full distribution of effect sizes (Kim *et al.* 2017) instead of a single s_d value for new deleterious mutations has no impact on the effect of *U* on $\frac{N_e}{N}$ (Figure 2C).

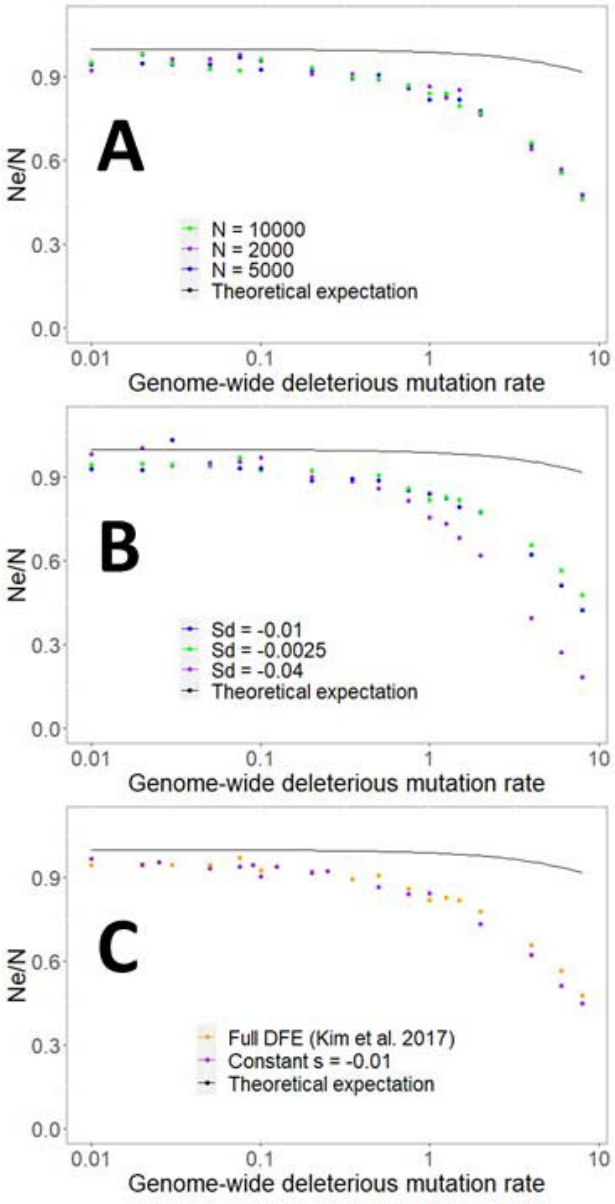

Figure 2. At high deleterious mutation rates, effective population sizes are much lower in multi-locus simulations than in analytic approximations which assume linkage equilibrium. Solid black line is the analytical expectation (Hudson and Kaplan 1995). Where not shown, $s_d = -0.01$ and N = 5000. (A) Census population size has no effect on $\frac{N_e}{N}$ ratio across a five-fold change in census population size. (B) Larger selective effects of deleterious mutations results in greater reduction of N_e at high mutation rates. (C) The relationship between deleterious mutation rate and N_e is the same whether the effect size of new deleterious mutations is constant versus. drawn from a full distribution of effect sizes (Kim *et al.* 2017) with the same mean value of -0.01. Each point represents a single simulation — we chose to allocate computation to a denser grid of parameter values rather than to replicates of the same parameter values.

The simulations above assume deleterious mutations occur uniformly at random across the genome. A more realistic scenario would be for deleterious mutations to be clustered within a functional subset of the genome. We modify our simulations to model genomes where only 10% of the genome is made up of 'genes' subject to deleterious mutations. Concentrating deleterious mutations into more tightly linked "genes" has mild and inconsistent effects (Figure 3, red vs. blue).

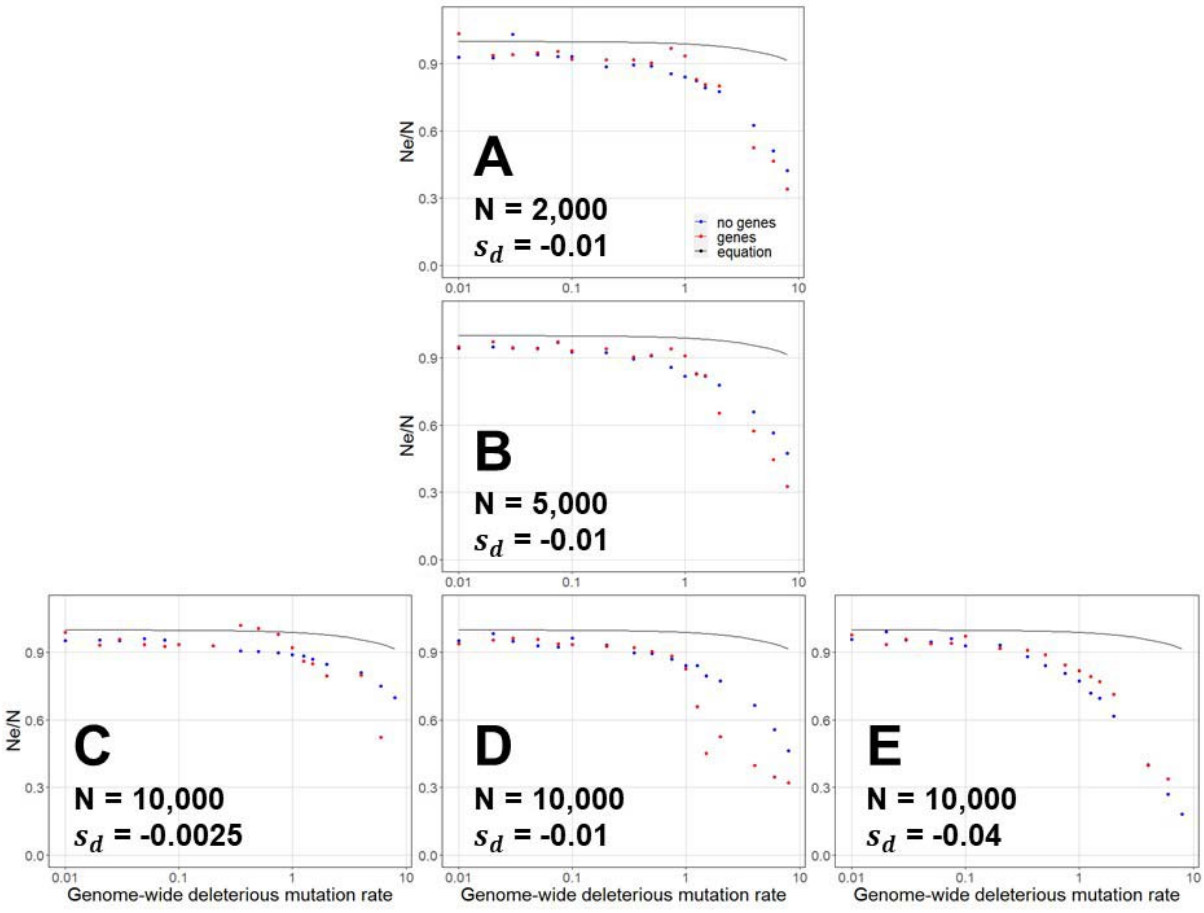

Figure 3. Concentrating deleterious mutations into "genes" has little effect on background selection. The strongest effect is seen at high population size, moderate strength of selection, and high mutation rate, where background selection clustered in genes reduces neutral diversity more than uniformly distributed background selection does. In all panels, the solid black line is the theoretical expectation from two-locus analytical approximations (Hudson and Kaplan 1995), blue dots are simulations where deleterious mutations occur uniformly at random on the genome, and red dots are simulations where deleterious mutations are clustered into 'genes'. Vertical comparison explores different census population sizes, and horizontal comparison explores different selective effects.

DISCUSSION

When deleterious mutation rates are realistically high, multi-locus simulations of evolution in the presence of background selection produce less neutral diversity than expected from analytic models that treat each deleterious mutation independently. This finding does not depend on the census population size, depends little on the selection coefficient characterizing sites under purifying selection, and depends little on the degree to which such sites are clustered near one another along chromosomes.

The disagreement between simulations and analytic models is substantial only with high deleterious mutation rates producing a significant number of tightly linked deleterious mutations subject to linkage disequilibrium. However, high deleterious mutation rates (Kondrashov and Crow 1993; Awadalla *et al.* 2010; Eöry *et al.* 2010; Lynch 2010; Roach *et al.* 2010; Kondrashov 2017) and widespread linkage disequilibrium (Conrad *et al.* 2006; Hinds *et al.* 2006; Koch *et al.* 2013) are both well established empirically. The effect of background selection on neutral diversity can thus be presumed to be larger than predicted by approximations that assume linkage equilibrium.

Our multilocus simulations neglect some population features known to affect neutral diversity (e,g, adaptive evolution (Maynard Smith and Haigh 1974) and temporal changes in population size (Torres *et al.* 2020)), and simplify others, (e.g. variation in dominance coefficients among deleterious variants (Gilbert *et al.* 2020) and heterogeneity in recombination rates (Kulathinal *et al.* 2008)). The purpose of our simulations is to isolate the effects of background selection with high mutation rates, rather than to accurately reflect the genetics of specific populations. Our simplified simulations nonetheless produce substantial disagreements with previous analytic methods, posing a serious challenge to the validity of those approximations. Incorporating additional complications into the model has more potential to strengthen rather than to weaken the broader case that simple analytic approximations are insufficient.

Debates about "neutral theory" have in recent years focused on whether patterns of genetic diversity can be explained by a combination of genetic drift, demography, and background selection (Lohmueller *et al.* 2011; Comeron 2014; Jensen *et al.* 2019), or whether these causes are insufficient and observed patterns indicate pervasive adaptation (Sella *et al.* 2009; Kern and Hahn 2018; Uricchio *et al.* 2019; Buffalo and Coop 2020). Our current work does not address this dispute about the relative importance of background selection vs. hitch-hiking. We instead exclude beneficial mutations in order to focus on models of background selection, extending them into the parameter regime of realistically high genome-wide deleterious mutation rates at which multi-locus effects can become significant.

However, our work does relate to whether it is appropriate to consider models of background selection to be a part of "neutral theory". While this might seem like a strange proposition, given that background selection obviously from its very name involves selection, proponents of "neutral theory" (Jensen *et al.* 2019) now include in their definition of the theory not only the direct effects of slightly deleterious mutations that were first treated by nearly neutral theory, but even the effects of that selection on linked neutral sites.

Behind these odd semantics is a substantial claim that background selection is a straightforward expansion of Kimura's original neutral theory. This claim is based on the argument that Kimura's original neutral models are still useful because the effects of background selection on neutral diversity can be captured by simply modifying the effective population size in a one-locus neutral model. Our results add to the body of evidence that this is not so simple. We find that background selection removes more neutral diversity than expected from previous two-locus models; this larger effect might broaden the scope of phenomena that background selection could explain. But we also find that there is no simple way to derive a value for an effective population size, which is highly sensitive to the genome-wide deleterious mutation rate U.

We can distinguish three criteria influencing judgements of any evolutionary theory, whether it be Kimura's original neutral theory, Ohta's nearly neutral theory, background selection theory, or hitch-

hiking theory. First, a theory's predictions must fit empirical patterns of genetic diversity — failure to meet this standard is what led to the replacement of Kimura's neutral theory by Ohta's nearly neutral theory.

Second, those predictions should be grounded in biologically reasonable and parsimonious assumptions. What these are can be subject to dispute, e.g. whether adaptive mutations should be excluded from "baseline models" (Comeron 2017; Johri *et al.* 2021) vs. whether independent lines of evidence so conclusively support widespread adaptation (Pennings *et al.* 2014; Enard *et al.* 2016) such that adaptive mutations must be considered until sufficient power for their exclusion has been proved. For our own somewhat different purposes, we take the empirical evidence for high deleterious mutation rates and widespread linkage disequilibrium to be broadly accepted, and hence their consideration to be required.

The third criterion is that it is wonderful when predictions can be derived from a simple mathematical framework such as Kimura's one-locus models. However, our results cast serious doubt on the validity of one-locus effective population size approximations of background selection. Instead, we suggest that given high deleterious mutation rates, models not just of hitchhiking but also of background selection need to incorporate less elegant multilocus effects of selection.

DATA AVAILABILITY

Simulation code written in python, and graphs produced with R. Scripts available on github at www.github.com/MaselLab/BackgroundSelection

ACKNOWLEDGEMENTS

We thank Bruce Walsh for pointing us to the analytical results on background selection that sparked this work, Ryan Gutenkunst and David Enard for helpful discussions, and Kevin Thornton for assistance with fwdpy simulations.

FUNDING

This work was supported by the John Templeton Foundation [62028].

CONFLICT OF INTEREST

The authors declare no conflict of interest.

BIBLIOGRAPHY

Awadalla P., J. Gauthier, R. A. Myers, F. Casals, F. F. Hamdan, *et al.*, 2010 Direct measure of the de novo mutation rate in autism and schizophrenia cohorts. Am. J. Hum. Genet. 87: 316–324. https://doi.org/10.1016/j.ajhg.2010.07.019

Barton N. H., 1998 The effect of hitch-hiking on neutral genealogies. Genet. Res. 72: 123–133. https://doi.org/10.1017/S0016672398003462

Barton N. H., and S. P. Otto, 2005 Evolution of recombination due to random drift. Genetics 169: 2353–2370. https://doi.org/10.1534/genetics.104.032821

Buffalo V., and G. Coop, 2020 Estimating the genome-wide contribution of selection to temporal allele frequency change. Proc. Natl. Acad. Sci. U. S. A. 117: 20672–20680. https://doi.org/10.1073/pnas.1919039117

Charlesworth B., M. T. Morgan, and D. Charlesworth, 1993 The Effect of Deleterious Mutations on Neutral Molecular Variation. Genetics 134: 1289–1303.

Charlesworth B., 2009 Effective population size and patterns of molecular evolution and variation. Nat. Rev. Genet. 10: 195–205. https://doi.org/10.1038/nrg2526

Comeron J. M., 2014 Background Selection as Baseline for Nucleotide Variation across the Drosophila Genome. PLoS Genet. 10: e1004434. https://doi.org/10.1371/journal.pgen.1004434

Comeron J. M., 2017 Background selection as null hypothesis in population genomics: Insights and

challenges from drosophila studies. Philos. Trans. R. Soc. B Biol. Sci. 372: 20160471. https://doi.org/10.1098/rstb.2016.0471

Conrad D. F., M. Jakobsson, G. Coop, X. Wen, J. D. Wall, *et al.*, 2006 A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. Nat. Genet. 38: 1251–1260. https://doi.org/10.1038/ng1911

Cutter A. D., and B. A. Payseur, 2013 Genomic signatures of selection at linked sites: unifying the disparity among species. Nat. Rev. Genet. 14: 262–274. https://doi.org/10.1038/nrg3425.Genomic

Enard D., L. Cai, C. Gwennap, and D. A. Petrov, 2016 Viruses are a dominant driver of protein adaptation in mammals. Elife 5: e12469. https://doi.org/10.7554/eLife.12469

Eöry L., D. L. Halligan, and P. D. Keightley, 2010 Distributions of selectively constrained sites and deleterious mutation rates in the hominid and murid genomes. Mol. Biol. Evol. 27: 177–192. https://doi.org/10.1093/molbev/msp219

Ewens W., 2004 *Mathematical Population Genetics*. Springer US, New York.

Felsenstein J., 1974 The evolutionary advantage of recombination. Genetics 78: 737–756. https://doi.org/10.1093/genetics/83.4.845

Gilbert K. J., F. Pouyet, L. Excoffier, and S. Peischl, 2020 Transition from Background Selection to Associative Overdominance Promotes Diversity in Regions of Low Recombination. Curr. Biol. 30: 101-107.e3. https://doi.org/10.1016/j.cub.2019.11.063

Haag-Liautard C., M. Dorris, X. Maside, S. Macaskill, D. L. Halligan, *et al.*, 2007 Direct estimation of per nucleotide and genomic deleterious mutation rates in Drosophila. Nature 445: 82–85. https://doi.org/10.1038/nature05388

Haller B. C., and P. W. Messer, 2017 SLiM 2: Flexible, interactive forward genetic simulations. Mol. Biol. Evol. 34: 230–240. https://doi.org/10.1093/molbev/msw211

Hill W. G., and A. Robertson, 1968 Linkage disequilibrium in finite populations. TAG Theor. Appl. Genet. 38: 226–231. https://doi.org/10.1007/bf01245622

Hinds D. A., A. P. Kloek, M. Jen, X. Chen, and K. A. Frazer, 2006 Common deletions and SNPs are in linkage disequilibrium in the human genome. Nat. Genet. 38: 82–85. https://doi.org/10.1038/ng1695

Hudson R. R., and N. L. Kaplan, 1994 Gene Trees with Background Selection, pp. 140–153 in *Non-neutral Evolution: Theories and Molecular Data*, edited by Golding B. Springer US, Boston, MA.

Hudson R. R., and N. L. Kaplan, 1995 Deleterious background selection with recombination. Genetics 141: 1605–1617.

Jensen J. D., B. A. Payseur, W. Stephan, C. F. Aquadro, M. Lynch, *et al.*, 2019 The importance of the Neutral Theory in 1968 and 50 years on: a response to Kern and Hahn 2018. Evolution (N. Y). 73: 111–114. https://doi.org/10.1111/evo.13650

Johri P., C. F. Aquadro, M. Beaumont, B. Charlesworth, L. Excoffier, *et al.*, 2021 Statistical Inference in Population Genomics. biorxiv 2021.10.27.466171. https://doi.org/10.1101/2021.10.27.466171

Keightley P. D., and S. P. Otto, 2006 Interference among deleterious mutations favours sex and recombination in finite populations. Nature 443: 89–92. https://doi.org/10.1038/nature05049

Kelleher J., A. M. Etheridge, and G. McVean, 2016 Efficient Coalescent Simulation and Genealogical Analysis for Large Sample Sizes. PLoS Comput. Biol. 12: e1004842. https://doi.org/10.1371/journal.pcbi.1004842

Kelleher J., K. R. Thornton, J. Ashander, and P. L. Ralph, 2018 Efficient pedigree recording for fast population genetics simulation. PLoS Comput. Biol. 14: e1006581.

Kern A. D., and M. W. Hahn, 2018 The Neutral Theory in Light of Natural Selection. Mol. Biol. Evol. 35: 1366–1371. https://doi.org/10.1093/molbev/msy092

Kim Y., and W. Stephan, 2000 Joint effects of genetic hitchhiking and background selection on neural variation. Genetics 155: 1415–1427. https://doi.org/10.1093/genetics/155.3.1415

Kim B., C. Huber, and K. Lohmueller, 2017 Inference of the Distribution of Selection Coefficients for New Nonsynonymous Mutations Using Large Samples. Genetics 206: 345–361. https://doi.org/10.1534/genetics.116.197145/-/DC1.1

Kimura M., 1968 Evolutionary Rate at the Molecular Level. Nature 217: 624–626.

Kimura M., 1969 The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. Genetics 61: 893–903.

King J. L., and T. H. Jukes, 1969 Non-Darwinian Evolution. Science (80-.). 164: 788–798.

Koch E., M. Ristroph, and M. Kirkpatrick, 2013 Long range linkage disequilibrium across the human genome. PLoS One 8: e80754. https://doi.org/10.1371/journal.pone.0080754

Kondrashov A. S., and J. F. Crow, 1993 A molecular approach to estimating the human deleterious mutation rate. Hum Mutat 2: 229–234. https://doi.org/10.1002/humu.1380020312

Kondrashov A. S., 2017 *Crumbling Genome: The Impact of Deleterious Mutations on Humans*. John Wiley & Sons, Inc, Hoboken, NJ.

Kulathinal R. J., S. M. Bennett, C. L. Fitzpatrick, and M. A. F. Noor, 2008 Fine-scale mapping of recombination rate in Drosophila refines its correlation to diversity and divergence. Proc. Natl. Acad. Sci. U. S. A. 105: 10051–10056. https://doi.org/10.1073/pnas.0801848105

Lesecque Y., P. D. Keightley, and A. Eyre-Walker, 2012 A resolution of the mutation load paradox in humans. Genetics 191: 1321–1330. https://doi.org/10.1534/genetics.112.140343

Lohmueller K. E., A. Albrechtsen, Y. Li, S. Y. Kim, T. Korneliussen, *et al.*, 2011 Natural Selection Affects Multiple Aspects of Genetic Variation at Putatively Neutral Sites across the Human

Genome. PLoS Genet. 7: e1002326. https://doi.org/10.1371/journal.pgen.1002326

Lynch M., 2010 Rate, molecular spectrum, and consequences of human mutation. Proc. Natl. Acad. Sci. U. S. A. 107: 961–8. https://doi.org/10.1073/pnas.0912629107

Masel J., 2011 Genetic drift. Curr. Biol. 21: R837–R838. https://doi.org/10.1016/j.cub.2011.08.007

Maynard Smith J., and J. Haigh, 1974 The hitch-hiking effect of a favourable gene. Genet. Res. (Camb). 23: 23–35. https://doi.org/10.1017/S0016672308009579

McEvoy B. P., J. E. Powell, M. E. Goddard, and P. M. Visscher, 2011 Human population dispersal "Out of Africa" estimated from linkage disequilibrium and allele frequencies of SNPs. Genome Res. 21: 821–829. https://doi.org/10.1101/gr.119636.110

Nordborg M., B. Charlesworth, and D. Charlesworth, 1996 The effect of recombination on background selection. Genet. Res. (Camb). 67: 159–174.

Ohta T., 1973 Slightly Deleterious Mutant Substitutions in Evolution. Nature 246: 96–98.

Pardo-Manuel De Villena F., and C. Sapienza, 2001 Recombination is proportional to the number of chromosome arms in mammals. Mamm. Genome 12: 318–322. https://doi.org/10.1007/s003350020005

Pennings P. S., S. Kryazhimskiy, and J. Wakeley, 2014 Loss and Recovery of Genetic Diversity in Adapting Populations of HIV. PLoS Genet. 10: e1004000. https://doi.org/10.1371/journal.pgen.1004000

Roach J. C., G. Glusman, A. F. A. Smit, C. D. Huff, P. T. Shannon, *et al.*, 2010 Analysis of Genetic Inheritance in a Family Quartet by Whole Genome Sequencing. Science (80-.). 328: 636–639. https://doi.org/10.1126/science.1186802.Analysis

Sella G., D. A. Petrov, M. Przeworski, and P. Andolfatto, 2009 Pervasive Natural Selection in the

Drosophila Genome? PLoS Genet. 5: e1000495. https://doi.org/10.1371/journal.pgen.1000495

Sung W., M. S. Ackerman, S. F. Miller, T. G. Doak, and M. Lynch, 2012 Drift-barrier hypothesis and mutation-rate evolution. Proc. Natl. Acad. Sci. 109: 18488–18492. https://doi.org/10.1073/pnas.1216223109/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1216223109

Takahata N., 1993 Allelic genealogy and human evolution. Mol. Biol. Evol. 10: 2–22. https://doi.org/10.1093/oxfordjournals.molbev.a039995

Tenesa A., P. Navarro, B. J. Hayes, D. L. Duffy, G. M. Clarke, *et al.*, 2007 Recent human effective population size estimated from linkage disequilibrium. Cold Spring Harb. Lab. Press Hum. 2: 520–526. https://doi.org/10.1101/gr.6023607.8

Thornton K. R., 2014 A c++ template library for efficient forward-time population genetic simulation of large populations. Genetics 198: 157–166. https://doi.org/10.1534/genetics.114.165019

Torres R., M. G. Stetter, R. D. Hernandez, and J. Ross-Ibarra, 2020 The Temporal Dynamics of Background Selection in Nonequilibrium Populations. Genetics 214: 1019–1030.

Uricchio L. H., D. A. Petrov, and D. Enard, 2019 Exploiting selection at linked sites to infer the rate and strength of adaptation. Nat. Ecol. Evol. 3: 977–984. https://doi.org/10.1038/s41559-019-0890-6

Vucetich J., T. Waite, and L. Nunney, 2017 Fluctuating Population Size and the Ratio of Effective to Census Population Size. Evolution (N. Y). 51: 2017–2021.

Wang J., and M. C. Whitlock, 2003 Estimating effective population size and migration rates from genetic samples over space and time. Genetics 163: 429–446. https://doi.org/10.1093/genetics/163.1.429

Wright S., 1938 Size of population and breeding structure in relation to evolution. Science (80-.). 87: 430–431.