doi:10.1111/evo.13650

The importance of the Neutral Theory in 1968 and 50 years on: A response to Kern and Hahn 2018

Jeffrey D. Jensen,[1,2] Bret A. Payseur,[3] Wolfgang Stephan,[4] Charles F. Aquadro,[5] Michael Lynch,[6] Deborah Charlesworth,[7] and Brian Charlesworth[7]

[1]*School of Life Sciences, Arizona State University, Tempe, Arizona*

[2]*E-mail: jeffrey.d.jensen@asu.edu*

[3]*Laboratory of Genetics, University of Wisconsin-Madison, Madison, Wisconsin*

[4]*Leibniz-Institute for Evolution and Biodiversity Science, Berlin, Germany*

[5]*Department of Molecular Biology & Genetics, Cornell University, Ithaca, New York*

[6]*Center for Mechanisms of Evolution, Arizona State University, Tempe, Arizona*

[7]*Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh, United Kingdom*

Received October 30, 2018

Accepted November 9, 2018

A recent article reassessing the Neutral Theory of Molecular Evolution claims that it is no longer as important as is widely believed. The authors argue that "the neutral theory was supported by unreliable theoretical and empirical evidence from the beginning, and that in light of modern, genome-scale data, we can firmly reject its universality." Claiming that "the neutral theory has been overwhelmingly rejected," they propose instead that natural selection is the major force shaping both between-species divergence and within-species variation. Although this is probably a minority view, it is important to evaluate such claims carefully in the context of current knowledge, as inaccuracies can sometimes morph into an accepted narrative for those not familiar with the underlying science. We here critically examine and ultimately reject Kern and Hahn's arguments and assessment, and instead propose that it is now abundantly clear that the foundational ideas presented five decades ago by Kimura and Ohta are indeed correct.

KEY WORDS: **Population Genetics, The Neutral Theory of Molecular Evolution.**

The Neutral Theory of Molecular Evolution asserts that most de novo mutations are either sufficiently deleterious in their effects on fitness that they have little chance of becoming fixed in the population, or are under such weak selection that they may become fixed as a result of genetic drift (Kimura 1968, 1983; King and Jukes 1969). Furthermore, the rate of substitution of neutral mutations between species is equal to the mutation rate (Kimura 1968). A critical first extension of this framework involved the inclusion of nearly neutral mutations, along with the recognition that the proportion of the genome represented by selectively constrained sites (where mutations have low probabilities of fixation by drift) depends on the effective population size of the species or genomic region (Ohta 1973). While drifting to fixation or loss, neutral and nearly neutral mutations contribute to DNA sequence variation within populations. The Neutral Theory further hypothesizes that advantageous mutations are sufficiently rare, compared to the constant input of neutral and deleterious variants, that they should be rarely present in samples of segregating variation, especially because of their rapid spread to fixation.

These ideas greatly changed the thinking of evolutionary biologists. Genetic drift was taken much more seriously than previously, stimulating a large body of fruitful empirical research into molecular evolution and variation, as well as fundamental advances in the stochastic theory of evolution, summarized in

Kimura's influential book (Kimura 1983). It is now difficult to appreciate how radical a departure this view of evolution represented: in the 1950s and 1960s, almost all evolutionary changes were attributed to directional natural selection, and most polymorphisms with alleles at intermediate frequencies were thought to be maintained by balancing selection (e.g., Ford 1975). Despite his pioneering contributions to stochastic population genetic theory, Fisher famously rejected any significant evolutionary role for genetic drift (Fisher 1930), though it is notable that Wright had simultaneously developed a deep appreciation for the importance of these stochastic effects that was later justified when molecular variants began to be studied (Wright 1931).

It is against this historical backdrop that Kern and Hahn (2018) discuss a purported controversy in population genetics concerning the predictive power and applicability of the Neutral Theory, beginning with the suggestion that "the ubiquity of adaptive variation both within and between species means that a more comprehensive theory of molecular evolution must be sought." Although those who initially developed the Neutral Theory did not claim that all sequence changes are neutral—indeed, Kimura himself developed some of the most fundamental theoretical formulations of selection and its interactions with genetic drift—Kern and Hahn (2018) argue that modern data have demolished the original evidence supporting the Neutral Theory. This is not a new claim. For example, Gillespie criticized some of the original arguments in favor of neutrality (e.g., Gillespie 1991), and nearly identical views were expressed in Hahn (2008). The novelty of the arguments of Kern and Hahn (2018) mainly lies in their emphasis on the effects of selection at linked sites on patterns of variation within genomes. Accordingly, we focus primarily on this aspect of their paper. As will become clear, a major problem with Kern and Hahn's views arise from their narrow definition of the Neutral Theory, which they summarize as follows: "differences between species are due to neutral substitutions (not adaptive evolution), and (. . . .) polymorphisms within species are not only neutral but also have dynamics dominated by mutation-drift equilibrium."

To support this narrow view, Kern and Hahn argue for pervasive effects of selection, relying heavily on a small number of population-genomic studies suggesting that as many as 50% of amino-acid replacement substitutions in *Drosophila* are adaptive (see, for example, the review by Sella et al. 2009), which they claim contradicts Kimura's (1968, 1983) and King and Jukes' (1969) assertion that most such substitutions are caused by genetic drift. Apart from the inherent uncertainty in these estimates (discussed by Fay 2011), it is misleading to use them to make the general claim that the Neutral Theory is insufficient to explain genome-wide patterns of variation and evolution; these inferred frequencies of adaptive substitutions mostly concern only the small fraction of the genome that codes for proteins (e.g., $<2\%$ of the human genome; see Lander et al. 2001). Kern and Hahn further overstate the pervasiveness of adaptive substitutions by highlighting studies in humans and plants that focus on the limited subset of genes that evolve rapidly. The circularity involved in ignoring the vast majority of neutral or nearly neutral substitutions across the genome, and then rejecting a significant role for neutrality, hardly justifies the need for the "selection theory of molecular evolution" advocated by Hahn (2008).

Second, with regard to the effects of selection on linked neutral or nearly neutral sites, Kern and Hahn (2018) emphasize the well-established positive correlation between recombination rates and levels of variation that has been observed in several species (Cutter and Payseur 2013). They begin with the very strong assertion that "these results imply that almost no loci are free from the effects of selection, in any organism." This broad claim is unjustified, given that there are relatively few species for which such data are available. Although this correlation (first documented in *Drosophila melanogaster* by Begun and Aquadro 1992) indeed suggests that selection reduces neutral variation at linked sites through the process of hitchhiking, the mutagenic effects of recombination itself may also contribute to this pattern (Pratto et al. 2014; Arbeithuber et al. 2015). Hitchhiking can involve both selective sweeps caused by the spread of favorable mutations (Maynard Smith and Haigh 1974), and the removal of neutral variants closely linked to deleterious mutations—background selection (Charlesworth et al. 1993; Charlesworth 2012). In an explicit comparison between models of widespread purifying selection on weakly deleterious alleles versus recurrent positive selection on beneficial alleles, Lohmueller et al. (2011) found a much better fit of the former to the observed pattern in humans (see also Pouyet et al. 2018), as did Comeron (2014) for *Drosophila*.

Importantly, observations from eukaryotic genomes, including humans and mice, show that levels of polymorphism are low in the neighborhood of coding or conserved noncoding sequences and increase approximately monotonically away from them (Cutter and Payseur 2013; Johri et al. 2017; Lynch et al. 2017). While selective sweeps may contribute to this pattern, and are indeed required to explain other observations (Campos et al. 2017), these findings imply that any selective sweeps involved must have rather local effects. Despite these results, Kern and Hahn (2018) emphasize studies that invoke pervasive positive selection to explain genome-wide patterns of variation (e.g., Garud et al. 2015; Schrider and Kern 2017). However, these claimed effects must be evaluated with caution owing to their failure to exclude or take proper account of the effects of the (unknown) non-equilibrium demographic histories of the populations in question.

Regardless of the precise interplay of the two forms of hitchhiking, background selection and selective sweeps, in shaping patterns of variation, it is important to note that neither affect the probability of fixation of neutral mutations (Birky and Walsh 1988), which determines the rate of neutral sequence evolution.

Both models are based on strong evidence that the vast majority of segregating variation is neutral or nearly neutral, and neither model contradicts the evidence that the vast majority of fixed differences between populations and species are also neutral or nearly neutral. Furthermore, both background selection and selective sweeps may be viewed as reducing the effective population size (N_e) of affected genomic regions, at least as a first approximation (see Charlesworth 2009). As shown by Kimura and Ohta (Kimura and Ohta 1971; Ohta 1973; Kimura 1983), a reduction in N_e causes the fixation probabilities of mutations with selective effects to be closer to those of neutral mutations, such that the rate of fixation of beneficial mutations is reduced, and the rate of fixation of deleterious mutations is increased—thereby, increasing the fraction of mutations that behave as effectively neutral. Thus, these hitchhiking effects only further emphasize the fundamental evolutionary role of genetic drift. Although the earliest formulations of the Neutral Theory focused on the dynamics of individual loci, and the effects of selection in reducing the N_e values at linked loci were not studied, we could not have understood these patterns without the contributions of Kimura and Ohta. It is simply a misunderstanding of the role of theoretical models in illuminating the interpretation of data to claim, as do Kern and Hahn (2018), that hitchhiking effects imply that levels of polymorphism are not at mutation-drift equilibrium, and "therefore, current data appear to be fundamentally incompatible with the neutral theory."

Given these considerations, we here propose a simple interpretation of the existing evidence in terms of a modern version of the Neutral Theory, whose individual components should not be controversial:

(1) A large fraction of the genome of organisms studied to date is subject to mutations that are effectively neutral with respect to their fitness effects, and hence evolve under genetic drift.
(2) The great majority of newly arising mutations that do affect fitness (i.e., non-neutral mutations) are deleterious, and the predominant mode of natural selection is purifying in nature, removing these deleterious mutations from populations.
(3) Natural populations are rarely at demographic equilibrium, and commonly have undergone recent historical changes. The combined effects of population size changes, structure, and migration all shape patterns of within-species variation. These demographic histories cannot be assumed to affect patterns of variation uniformly across the genome, and indeed may produce different effects in different genomic regions, mimicking expectations under selection (e.g., Wall et al. 2002; Thornton and Jensen 2007).
(4) A combination of genetic drift (as modulated by the demographic history of the population) with both direct and linked purifying selection shapes patterns of genomic variation. Thus, a model taking joint account of all of these effects is essential for genomic analysis (Comeron 2017), and progress is being made towards this goal (e.g., Zeng and Charlesworth 2010).
(5) Beneficial mutations occasionally arise and some may reach fixation or high frequencies, and localized hitchhiking effects related to such events have been convincingly described in a variety of organisms. In some cases, these genotypic changes have been meaningfully connected with both phenotype and fitness. However, the effects of these comparatively rare, localized positive selection events are best characterized and quantified as additional to the genome-wide processes described above (Stephan 2010). In the absence of an appropriate null model accounting for these processes that are common to the genome as a whole, inappropriate adaptive story-telling will be likely to proliferate.

All five points are fully consistent with the ground-breaking work of Kimura and Ohta. Furthermore, developments made in the light of empirical observations subsequent to Kimura's initial publication are straightforward extensions of the Neutral Theory. They demonstrate its continued importance, rather than demolishing it. Over the past five decades, such insights have enhanced our understanding of the interplay of population size with drift-selection dynamics (Ohta 1973), and described the hitchhiking effects of selection induced by the comparatively rare class of beneficial mutations (Maynard Smith and Haigh 1974), as well as those caused by the much more common class of deleterious mutations (Charlesworth et al. 1993). This framework has also served as an organizing principle for understanding patterns of variation in genome architecture (Lynch 2007), and for understanding the evolution of cellular features, including the mutation rate itself (Lynch et al. 2016).

Thus, our use of the term "ground-breaking" to describe the Neutral Theory is not meant to imply a scientific advance that was fully formed at the outset. Like other major scientific advances, the Neutral Theory has been adjusted and modified over time in light of later observations and thought, yet retains its value. For example, Darwin's findings and reasoning supporting the operation of natural selection were not abandoned owing to his lack of a satisfactory theory of heredity—indeed, the incorporation of that subsequent knowledge only strengthened the underlying concepts (Fisher 1930). Similarly, the Neutral Theory should not be dismissed because of the lack of emphasis on the effects of selection at linked sites in its initial formulation, as subsequent studies have only served to emphasize the fundamental role of near neutrality and genetic drift in shaping the variation observed within and between species. Indeed, Ohta and Kimura were among the first to study such effects, in their analysis of the apparent overdominance at neutral sites induced by linkage to sites subject to heterozygote advantage or selection against deleterious mutations (Ohta and Kimura 1970; Ohta 1971).

In sum, the transition to molecular biology has increased the importance of population genetics for our understanding of evolution. Moreover, instead of unraveling the prior theoretical framework, the influx of molecular data has lent support to many pre-genomic theoretical developments. Although the edifice may not yet be complete, the Neutral Theory changed how people thought about evolution at the molecular level, and this framework appropriately continues to serve as the basis of modern evolutionary genomics. Thus, great credit is owed to the scientists who worked this theory out in detail and anticipated much of what it could tell us once genes (and genomes) could be sequenced.

ACKNOWLEDGMENTS

The authors thank Peter Andolfatto, Andy Clark, Laurent Excoffier, Adam Eyre-Walker, Peter Keightley, Michael Nachman, Rasmus Nielsen, Mohamed Noor, Molly Przeworski, Kevin Thornton, John Wakeley, John Willis, and members of the Jensen Lab for supportive comments and helpful suggestions.

LITERATURE CITED

Arbeithuber, B., A. J. Betancourt, T. Ebner, and I. Tiemann-Boege. 2015. Crossovers are associated with mutation and biased gene conversion at recombination hotspots. Proc. Natl. Acad. Sci. USA 112:2109–14.

Begun, D. J. and C. F. Aquadro. 1992. Levels of naturally occurring DNA polymorphism correlated with recombination rates in *D. melanogaster*. Nature 356:519–20.

Birky, C. W., and J. B. Walsh. 1988. Effects of linkage on rates of molecular evolution. Proc. Natl. Acad. Sci. USA 85:6414–18.

Campos, J. L., L. Zhao, and B. Charlesworth. 2017. Estimating the parameters of background selection and selective sweeps in Drosophila in the presence of gene conversion. Proc. Natl. Acad. Sci. USA 114:E4762–71.

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(3):195–205.

———. 2012. The effects of deleterious mutations on evolution at linked sites. Genetics 190:5–22.

Comeron, J. M. 2014. Background selection as baseline for nucleotide variation across the Drosophila genome. PLoS Genet. 10(6):e1004434.

———. 2017. Background selection as null hypothesis in population genomics: insights and challenges from Drosophila studies. Phil. Trans. R. Soc. B 372(1736).

Cutter, A. D. and B. A. Payseur. 2013. Genomics signatures of selection at linked sites: unifying the disparity among species. Nat. Rev. Genet. 14(4):262–74.

Fay, J. C. 2011. Weighing the evidence for adaptation at the molecular level. Trends Genet. 27(9):343–9.

Fisher, R. A. 1930. The genetical theory of natural selection. Clarendon Press, Oxford.

Ford, E. B. 1975. Ecological genetics. Chapman and Hall, London.

Garud, N., P. Messer, E. Buzbas, and D. Petrov. 2015. Recent selective sweeps in North American *Drosophila melanogaster* show signatures of soft sweeps. PLoS Genet. 11:e1005004.

Gillespie, J. H. 1991. The causes of molecular evolution. Oxford Univ. Press, Oxford, England.

Hahn, M. W. 2008. Toward a selection theory of molecular evolution. Evolution 62:255–65.

Johri, P., S. Krenek, G. K. Marinov, T. G. Doak, T. U. Berendonk, and M. Lynch. 2017. Population genomics of *Paramecium* species. Mol. Biol. Evol. 34:1194–1216.

Kern, A. D. and M. W. Hahn. 2018. The neutral theory in light of natural selection. Mol. Biol. Evol. 35:1366–71.

Kimura, M. 1968. Evolutionary rate at the molecular level. Nature 217:624–6.

———. 1983. The neutral theory of molecular evolution. Cambridge Univ. Press, Cambridge, MA.

Kimura, M., and T. Ohta. 1971. Theoretical aspects of population genetics. Princeton Univ. Press, Princeton, England.

King, J. L., and T. H. Jukes. 1969. Non-Darwinian evolution. Science 164: 788–98.

Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. Zody, J. Baldwin, et al. 2001. Initial sequencing and analysis of the human genome. Nature 409:860–921.

Lohmueller, K. E., A. Albrechtsen, Y. Li, S. Kim, T. Korneliussen, N. Vinckenbosch, G. Tian, E. Huerta-Sanchez, A. Feder, N. Grarup, et al. 2011. Natural selection affects multiple aspects of genetic variation at putatively neutral sites across the human genome. PLoS Genet. 7:e1002326.

Lynch, M. 2007. The origins of genome architecture. Sinauer, Sunderland, MA.

Lynch, M., M. Ackerman, J.-F. Gout, H. Long, W. Sung, W. K. Thomas, and P. L. Foster. 2016. Genetic drift, selection, and the evolution of the mutation rate. Nat. Rev. Genet. 17:704–714.

Lynch, M., R. Gutenkunst, M. Ackerman, K. Spitze, Z. Ye, T. Maruki, and Z. Jia. 2017. Population genomics of *Daphnia pulex*. Genetics 206:315–32.

Maynard Smith, J., and J. Haigh. 1974. The hitch-hiking effect of a favourable gene. Genet. Res. 23:23–35.

Ohta, T., and M. Kimura. 1970. Development of associative overdominance through linkage disequilibrium in finite populations. Genet. Res. 16:165–77.

Ohta, T. 1971. Associate overdominance caused by linked detrimental mutations. Genet. Res. 17:277–86.

———. 1973. Slightly deleterious mutant substitutions in evolution. Nature 246:96–8.

Pouyet, F., S. Aeschbacher, A. Thiery, and L. Excoffier. 2018. Background selection and biased gene conversion affect more than 95% of the human genome and bias demographic inference. ELife 7:e36317.

Pratto, F., K. Brick, P. Khil, F. Smagulova, G. V. Petukhova, and R. D. Camerini-Otero. 2014. DNA Recombination. Recombination initiation maps of individual human genomes. Science 346:1256442.

Schrider, D. R. and A. D. Kern. 2017. Soft sweeps are the dominant mode of adaptation in the human genome. Mol. Biol. Evol. 34:1863–77.

Sella, G., D. Petrov, M. Przeworski, and P. Andolfatto. 2009. Pervasive natural selection in the Drosophila genome? PLoS Genet. 5:e1000495.

Stephan, W. 2010. Genetic hitchhiking versus background selection: the controversy and its implications. Phil. Trans. R. Soc. B 365:1245–53.

Thornton, K. R. and J. D. Jensen. 2007. Controlling the false-positive rate in multi-locus genome scans for selection. Genetics 175:737–50.

Wall, J. D., P. Andolfatto, and M. Przeworski. 2002. Testing models of selection and demography in *Drosophila simulans*. Genetics 162:203–16.

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16:97–159.

Zeng, K. and B. Charlesworth. 2010. The effects of demography and linkage on the estimation of selection and mutation parameters. Genetics 186: 1411–24.

Handling Editor: Mohamed A. F. Noor
Associate Editor: Mohamed A. F. Noor