**Falsifying the Kimura Fixation Model:**

**The Ne Equivocation and the Empirical Failure of Neutral Theory**

Vox Day and Claude Athos

Correspondence: voxday@gmail.com

**Abstract**

Kimura's neutral theory of molecular evolution rests on two foundational results: the substitution rate *k* = μ and the equilibrium diversity π = 4*N*eμ. Both formulas employ effective population size (*N*e), but we demonstrate that *N*e cannot represent the same biological quantity in both contexts. The substitution rate derivation requires *N*e to govern both mutation supply and fixation probability; however, mutations arise from the census population while fixation probability depends on drift dynamics. Using empirical data from three mammalian species (chimpanzees, gorillas, and killer whales), we show that census populations exceed diversity-derived *N*e by 19- to 46-fold. This systematic discrepancy reveals that *N*e functions as a post-hoc fitting parameter rather than an independent measurement. We further demonstrate that explaining these *N*e reductions requires invoking pervasive selection. Combined with empirical falsification from ancient DNA time series (Bio-Cycle model reduces Kimura prediction error by 69%), we conclude that the Kimura model is degenerately parameterized and does not describe the evolutionary dynamics of sexually reproducing organisms with overlapping generations.

**Keywords:** effective population size, neutral theory, molecular evolution, genetic drift, nucleotide diversity, Kimura model

# 1. Introduction

Motoo Kimura's neutral theory of molecular evolution (Kimura 1968, 1983) represents one of the most influential frameworks in population genetics. The theory proposes that most genetic variation within species and most fixed differences between species are selectively neutral, arising through random genetic drift rather than natural selection. Two mathematical results form the foundation of neutral theory: (1) the rate of neutral substitution equals the mutation rate, independent of population size; and (2) equilibrium nucleotide diversity is proportional to effective population size times mutation rate.

These results have been enormously productive for molecular evolutionary analysis. The molecular clock, calibrated using fossil divergence times, depends on the constancy of neutral substitution rates. Coalescent theory, which reconstructs demographic history from genetic data, relies on the relationship between diversity and effective population size. However, the mathematical elegance of these results obscures a fundamental problem: the parameter *N*e appears in both formulas but cannot represent the same biological quantity in both contexts.

In this paper, we demonstrate that Kimura's neutral theory contains an internal contradiction regarding the meaning of effective population size, and we show that this contradiction has empirical consequences visible across multiple mammalian species. We further demonstrate that ancient DNA time series falsify the Kimura model's predictions, with an alternative model incorporating generation overlap (the Bio-Cycle model) reducing prediction error by 69%.

# 2. The Two-*N*e Problem

Before examining the technical argument, we distinguish three claims often bundled under “neutral theory”: (1) the rate of neutral substitution equals the mutation rate, independent of population size; (2) neutral diversity scales linearly with effective population size; (3) the boundary between neutral and selected alleles is governed by *N*e × *s* ≈ 1. While conceptually related, these claims have different empirical signatures and different vulnerabilities. Our critique of the Two-*N*e problem attacks claim (1) directly; our empirical tests address claims (1) and (2); the census-*N*e discrepancy undermines all three.

## 2.1 The Substitution Rate Derivation

Kimura's celebrated result that the neutral substitution rate equals the mutation rate derives from the cancellation of population size terms. In a diploid population of *N*e individuals, 2*N*eμ new neutral mutations arise per generation at each site. Each mutation has fixation probability 1/(2*N*e). The substitution rate is therefore:

*k* = 2*N*eμ × 1/(2*N*e) = μ

The *N*e terms cancel, yielding a substitution rate independent of population size. This mathematical elegance has been taken as evidence for the power of neutral theory.

## 2.2 The Diversity Prediction

The second foundational result predicts nucleotide diversity at equilibrium:

π = 4*N*eμ

This predicts that larger populations maintain more standing genetic variation because drift eliminates alleles more slowly. The formula provides the primary empirical method for estimating *N*e: given observed diversity π and an estimated mutation rate μ, one solves for *N*e = π/(4μ).

## 2.3 The Hidden Contradiction

Both formulas employ *N*e, but the substitution rate derivation requires *N*e to represent two distinct biological processes:

**Mutation supply:** The term 2*N*eμ represents the number of new mutations arising per generation. This depends on how many individuals are replicating DNA.

**Fixation probability:** The term 1/(2*N*e) represents the probability that a neutral mutation eventually reaches fixation. This depends on the effective dynamics of genetic drift.

For the cancellation to work, these two *N*e values must be equal. But mutation is a biochemical process occurring during DNA replication. Mutation origination occurs in census individuals, but the probability that a mutation enters the segregating population is weighted by reproductive success. Using a single *N*e to represent both instantaneous mutation origination and long-term lineage survival collapses distinct temporal processes into a single scalar.

Fixation probability, by contrast, depends on the dynamics of drift, which is governed by variance in reproductive success, population structure, and other factors captured in *N*e. The effective population size is typically much smaller than census size—often by an order of magnitude or more.

If census *N* governs mutation supply while *N*e governs fixation probability, the cancellation fails:

*k* = 2*N*μ × 1/(2*N*e) = (*N*/*N*e)μ

When *N* >> *N*e, the substitution rate exceeds the mutation rate by the ratio of census to effective population size.

# 3. Empirical Test: The Census-*N*e Discrepancy

If the Kimura model correctly describes real populations, we should find that census population size and diversity-derived *N*e are reasonably close—perhaps differing by a small factor reflecting known sources of reproductive variance. Instead, empirical data reveal systematic discrepancies of one to two orders of magnitude.

## 3.1 Chimpanzees (*Pan troglodytes*)

Census population estimates for chimpanzees in the early 20th century range from 1–2 million individuals (WWF 2012; various conservation sources). Current populations of 170,000–300,000 reflect habitat destruction and hunting pressure, not historical levels.

Yu et al. (2003) measured nucleotide diversity in chimpanzees using 50 randomly chosen noncoding DNA segments. They found π = 0.092–0.132% depending on subspecies, with all chimpanzees combined showing π = 0.132%. These values are only 1–1.5× higher than humans (π = 0.088%).

Using the standard formula with μ = 1.5 × 10-8:

*N*e = π/(4μ) = 0.0013/(6 × 10-8) ≈ 21,700

This matches the literature estimate precisely (Yu et al. 2003 report *N*e ≈ 21,300). But census *N* ≈ 1,000,000 exceeds this *N*e by a factor of **46**.

If we use census *N* for diversity prediction: π = 4 × 1,000,000 × 1.5 × 10-8 = 6%. The observed value of 0.13% is 46× lower than predicted.

## 3.2 Gorillas (*Gorilla gorilla*)

Western lowland gorilla census population is approximately 316,000 (IUCN Red List), with historical populations likely exceeding 500,000. Yu et al. (2004) measured π = 0.158% for western lowland gorillas—the highest among African apes but only about twice the human value.

Back-calculated *N*e: 0.00158/(6 × 10-8) ≈ 26,300. Yu et al. (2004) reported *N*e ≈ 25,200. Census *N* ≈ 500,000 exceeds this by a factor of **19**.

Eastern gorillas present an anomaly: with census *N* ≈ 5,000 and π = 0.055–0.08%, the back-calculated *N*e (9,000–13,000) *exceeds* census size—indicating that *N*e captures historical rather than current population dynamics.

## 3.3 Killer Whales (*Orcinus orca*)

Global killer whale census population is approximately 50,000–90,000 individuals. Hoelzel et al. (2002) documented remarkably low genetic diversity—virtually no mtDNA variation within some regional populations. This finding—"Low worldwide genetic diversity in the killer whale"—has prompted explanations involving matrilineal expansion and possible historical bottlenecks.

The discrepancy between census abundance and genetic diversity in killer whales exceeds that in great apes, with *N*e/*N* ratios potentially below 0.01.

## 3.4 Summary of Discrepancies

**Table 1.** Census population vs. diversity-derived *N*e in three mammalian species

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Species** | **Census N** | **Observed π** | **Diversity Ne** | **N/Ne Ratio** |
| Chimpanzee | ~1,000,000 | 0.13% | ~21,700 | 46:1 |
| Gorilla | ~500,000 | 0.16% | ~26,300 | 19:1 |
| Killer whale | ~50,000+ | Very low | Very low | >10:1 |

In every case, census *N* dramatically exceeds the *N*e required to match observed diversity. The discrepancies are not marginal factors of 2–3 that might be explained by reproductive variance; they are order-of-magnitude differences requiring explanation.

# 4. The Circular Defense

The standard response to census/*N*e discrepancies is that *N*e "captures all the deviations from Wright-Fisher assumptions." This response, while technically accurate, reveals that *N*e functions as a fitting parameter rather than an independent measurement.

## 4.1 What Reduces *N*e?

The mechanisms invoked to explain why *N*e << *N* include:

**Historical bottlenecks:** Past population crashes reduce diversity, which is slow to recover. But chimpanzee populations have been large for millions of years.

**Population structure:** Subdivision reduces effective size. But structured populations violate the random-mating assumption central to Wright-Fisher models.

**Background selection:** Deleterious mutations reduce diversity at linked neutral sites. But this is *selection*.

**Selective sweeps:** Beneficial mutations hitchhike linked variants to fixation, reducing diversity. This is also *selection*.

**Reproductive skew:** Unequal contribution to the next generation reduces effective size. But differential reproduction based on genotype *is* selection.

Every mechanism that reduces *N*e below census *N* by 20–50 fold either violates Wright-Fisher assumptions or invokes selection. A 46-fold reduction in chimpanzees implies that *selection-like processes* dominate genetic dynamics across the genome—yet neutral theory claims most of the genome evolves neutrally.

## 4.2 The Circularity

The effective population size is defined as "the size of an idealized Wright-Fisher population that would experience the same rate of genetic drift as the actual population" (Crow & Kimura 1970). In practice, *N*e is calculated from observed diversity:

*N*e = π/(4μ)

This means *N*e is *defined* as whatever value makes the diversity formula true. When we then use *N*e to “predict” diversity, we are engaged in circularity: the prediction reduces to the statement π = π.

We do not claim the model is logically unfalsifiable in the strict Popperian sense. One could imagine independent *N*e measurements from pedigree variance or linkage disequilibrium. But such measurements routinely disagree with diversity-based *N*e, which is then explained by invoking additional parameters (background selection, selective sweeps, population structure). The model is *degenerately parameterized*: it has more degrees of freedom than data points to constrain them. Any observed diversity can be accommodated by adjusting *N*e, leaving the theory without independent predictive content.

# 5. Lewontin's Paradox: The Field Knows

The disconnect between census population size and genetic diversity has been recognized since Lewontin (1974) observed that diversity across species varies only ~10-fold despite census sizes varying over many orders of magnitude. This “Paradox of Variation” challenges neutral theory, which predicts diversity proportional to an effective population size often interpreted as reflecting neutral drift. Explaining why *N*e is orders of magnitude smaller than census size requires invoking selective or structural processes that undermine the neutrality assumption.

Buffalo (2021) revisited this paradox using modern genomic data: "Under neutral theory, one would expect that pairwise diversity also vary over several orders of magnitude. However, early allozyme surveys revealed that diversity levels across a wide range of species varied just an order of magnitude... this is known as Lewontin's Paradox of Variation."

Proposed resolutions include non-equilibrium demography, variance in reproductive success, and linked selection. Notably, linked selection models (background selection and recurrent hitchhiking) invoke *selection* to explain patterns that neutral theory was supposed to explain *without* selection. This tension has been recognized since the 1970s; the question is whether linked selection can resolve it.

# 6. Empirical Falsification: The Bio-Cycle Model

The preceding arguments demonstrate internal inconsistency and degenerate parameterization with respect to diversity. A stronger test comes from ancient DNA time series, which allow direct observation of allele frequency trajectories.

Day and Athos (2025) introduced the Bio-Cycle Fixation Model, which incorporates a generation overlap correction factor (*d*) into standard allele frequency dynamics. Classical models assume discrete, non-overlapping generations (*d* = 1); the Bio-Cycle model allows *d* < 1 to account for the fact that multiple cohorts coexist and compete simultaneously in real populations.

The loci we examine are under positive selection and thus fall outside neutral theory’s core domain. However, Kimura’s fixation probability formula (and its dynamical extension) is routinely applied to selected alleles; our test evaluates whether the standard Wright-Fisher generation structure accurately models allele frequency change regardless of selective regime.

## 6.1 Ancient DNA Test

Three loci with well-documented ancient DNA time series and published selection coefficient estimates were analyzed:

**Lactase persistence (LCT/MCM6):** Rose from <1% to ~75% in Northern Europeans over ~6,000 years; *s* = 0.04–0.10

**SLC45A2:** Skin pigmentation locus; rose from ~43% to ~97% over ~4,000 years; *s* = 0.04–0.05

**TYR:** Secondary pigmentation locus; rose from ~25% to ~76% over ~5,000 years; *s* = 0.02–0.04

## 6.2 Results

Using published selection coefficients, the Bio-Cycle model (*d* = 0.45) substantially outperformed the Kimura model (*d* = 1.0) for all three loci:

**Table 2.** Prediction accuracy comparison

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Locus** | **Observed** | **Kimura pred.** | **Bio-Cycle pred.** | **Error reduction** |
| Lactase | 75% | 99.9% | 66.2% | 65% |
| SLC45A2 | 97% | 99.9% | 96.2% | 72% |
| TYR | 76% | 99.3% | 82.7% | 71% |

The classical Kimura model systematically over-predicts allele frequencies, driving alleles to near-fixation when observed frequencies remain substantially lower. The average error reduction of 69% demonstrates that the Bio-Cycle correction captures genuine population dynamics.

Cross-validation using the ratio of required selection coefficients at TYR and SLC45A2 (both under selection for the same phenotype) confirms that *d* ≈ 0.45 reflects demographic reality: the ratio remains constant at ~0.49 across all *d* values, as expected for genes responding to the same selective pressure with different effect sizes.

## 6.3 Theoretical Grounding

The empirical *d* ≈ 0.45 is consistent with theoretical predictions from the Selective Turnover Coefficient (Day & Athos 2025b), which derives *d* from reproductive-value-weighted mortality schedules using Coale-Demeny life tables. For Neolithic/Bronze Age demographic conditions, theory predicts *d* ≈ 0.45–0.53—bracketing the empirical estimate.

This consilience—an empirical discovery subsequently grounded in demographic theory—strengthens confidence in both the finding and its interpretation: effective generation counts in overlapping-generation species are approximately half the nominal counts.

# 7. Discussion

## 7.1 The Status of Kimura's Model

We have demonstrated three independent failures of the Kimura neutral theory:

**Internal contradiction:** The *N*e in the substitution rate derivation must represent both mutation supply (census *N*) and fixation probability (drift-effective *N*e). These differ by 20–50× in large mammals.

**Degenerate parameterization:** *N*e is calculated from diversity, then used to "predict" diversity. The model cannot fail because the free parameter absorbs all discrepancies.

**Empirical falsification:** Ancient DNA time series show systematic over-prediction by the Kimura model, with the Bio-Cycle correction reducing error by 69%.

Any one of these would be grounds for serious concern. Together, they demonstrate that the Kimura model does not describe the evolutionary dynamics of real populations.

## 7.2 What *N*e Actually Measures

The effective population size is often described as a "nuisance parameter" that absorbs all deviations from idealized assumptions. More precisely, *N*e measures "whatever happened that affected diversity"—including selection, population structure, and demographic history. It is not an independent measurement of population size; it is a back-calculation from genetic patterns.

Waples (2025) acknowledges this in “The Idiot’s Guide to Effective Population Size”: “Today, the effective population size is usually estimated empirically with respect to the amount of within-species genetic diversity divided by the mutation rate, yielding a coalescent effective population size that reflects the cumulative effects of genetic drift, *background selection*, and *genetic hitchhiking* over longer time periods" (emphasis added).

Note that background selection and genetic hitchhiking are both *selection*. The parameter *N*e absorbs selective effects that neutral theory was designed to avoid modeling.

## 7.3 Implications

The failure of the Kimura model does not invalidate all of population genetics. The mathematical framework remains useful for statistical inference and data analysis. But the model does not describe mechanism; it describes pattern.

Claims about evolutionary rates, fixation times, and the power of drift versus selection that depend on *N*e should be interpreted with caution. The parameter is not measuring what biologists often assume it measures.

For species with overlapping generations—which is to say, nearly all species of interest—the Bio-Cycle correction suggests that effective generation counts are approximately half the nominal counts. This has profound implications for estimates of evolutionary rates over deep time.

# 8. Conclusion

Kimura's neutral theory is elegant mathematics built on an unstable foundation. The *k* = μ result requires that mutation supply and fixation probability use the same *N*e, but these represent different biological processes governed by different population sizes. The diversity prediction is operationally unfalsifiable because *N*e is back-calculated from diversity. Ancient DNA falsifies the model's predictions with 69% error reduction under an alternative framework.

The model’s continued use reflects its flexibility rather than its predictive accuracy; the free parameter accommodates observations post-hoc. The theory is protected from falsification by auxiliary hypotheses that lack independent support.

Population genetics would benefit from acknowledging the persistent empirical pattern: nucleotide diversity does not scale with census population size as neutral theory predicts. Continued reliance on *N*e as a meaningful biological parameter obscures rather than illuminates evolutionary dynamics.

# References

Buffalo V (2021) Quantifying the relationship between genetic diversity and population size suggests natural selection cannot explain Lewontin's Paradox. *eLife* 10:e67509.

Crow JF, Kimura M (1970) *An Introduction to Population Genetics Theory*. Harper & Row, New York.

Day V, Athos C (2025a) The Bio-Cycle Fixation Model: Empirical Validation of a Generation Overlap Correction for Allele Frequency Dynamics. *Research Square* preprint.

Day V, Athos C (2025b) The Selective Turnover Coefficient: A First-Principles Derivation. *Research Square* preprint.

Hoelzel AR, Natoli A, Dahlheim ME, Olavarria C, Baird RW, Black NA (2002) Low worldwide genetic diversity in the killer whale (*Orcinus orca*): implications for demographic history. *Proceedings of the Royal Society B* 269:1467–1473.

Kimura M (1968) Evolutionary rate at the molecular level. *Nature* 217:624–626.

Kimura M (1983) *The Neutral Theory of Molecular Evolution*. Cambridge University Press.

Lewontin RC (1974) *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York.

Prado-Martinez J et al. (2013) Great ape genetic diversity and population history. *Nature* 499:471–475.

Waples RS (2025) The Idiot's Guide to Effective Population Size. *Molecular Ecology* doi:10.1111/mec.17670.

Yu N, Jensen-Seaman MI, Chemnick L, Kidd JR, Deinard AS, Ryder O, Kidd KK, Li W-H (2003) Low nucleotide diversity in chimpanzees and bonobos. *Genetics* 164:1511–1518.

Yu N, Jensen-Seaman MI, Chemnick L, Ryder O, Li W-H (2004) Nucleotide diversity in gorillas. *Genetics* 166:1375–1383.