Genomic data confirms that mutation cannot restore genetic diversity lost through population bottlenecks

Toby G. L. Kovacs[1], Simon Y. W. Ho[1], Carolyn J. Hogg[1] & Catherine E. Grueber[1]

[1]School of Life and Environmental Sciences, The University of Sydney, Sydney, New South Wales, Australia

Corresponding Author: Toby G. L. Kovacs, toby.kovacs@sydney.edu.au, (+61) 431 775 796

Emails of Co-authors: simon.ho@sydney.edu.au, carolyn.hogg@sydney.edu.au and catherine.grueber@sydney.edu.au

Manuscript Information:

- Type of article: *Perspective*
- Word count: 2849 (plus a 198-word abstract)
- Number of references: 33
- Figures: 2
- Tables: 0

Key Words

allelic richness, conservation, genetic diversity, heterozygosity, koala, mutation rate, population bottleneck, recovery, site frequency spectrum

ABSTRACT

Maintaining within-species genetic diversity is a critical goal of biodiversity conservation as it determines a species' ability to adapt to environmental change. Without human intervention, isolated populations can only recover genetic diversity post-bottleneck through the accumulation of new mutations over evolutionary timescales. Using recent estimates of mutation rates from major genomic datasets across the tree of life, we modelled the recovery of genetic diversity post-bottleneck. We found that mutation rates do not affect the recovery rate of genetic diversity, which instead is determined by the effective population size (N_e). For example, regardless of the mutation rate, an isolated population with $N_e = 500$ that experiences a 50% reduction in heterozygosity would require ~2,300 generations to return to 95% of its pre-disturbance level. In contrast, a bottleneck can lead to a 50% reduction in heterozygosity very quickly, taking just 30 generations for bottlenecks of $N_e = 20$. We also demonstrate that allelic richness responds quickly following the recovery of N_e but argue that this is unlikely to correspond to the recovery of evolutionary potential. Our results reinforce that recovery via mutation alone is too slow to be effective within conservation timeframes, providing an evolutionary context to genetic diversity loss.

1 Determinants of genetic diversity

Maintaining global biodiversity is essential for the regular and reliable function of the world's agricultural, economic and health systems (Scheffers et al. 2016). Despite this, anthropogenic influences continue to deteriorate ecosystems, leading to widespread losses in species (Cowie et al. 2022) and genetic diversity (Shaw et al. 2025). Fortunately, the importance of maintaining genetic diversity is becoming increasingly recognised in global conservation policy. The United Nations Convention on Biological Diversity (CBD) Kunming-Montreal Global Biodiversity Framework (GBF) marks a significant improvement in this regard, introducing its support for the preservation of genetic diversity of wild animals (CBD 2022)**.** This includes Target 4 of the agreement that aims to *maintain* and *restore* genetic diversity in order to maintain adaptive potential. Maintaining genetic diversity has long been the key focus of conservation genetic projects and there are a suite of tools available to management practitioners (Shaw et al. 2025). However, options available for restoring genetic diversity in wild populations are much more limited.

During a population bottleneck, genetic drift erodes genetic diversity, leading to the loss of evolutionary potential (Franklin 1980). This lost genetic variation can be summarised in two components: the number of segregating sites in a population (quantified as allelic richness or nucleotide diversity), and the evenness of their frequencies (quantified by expected heterozygosity, also known as gene diversity) (Hoban et al. 2022). These measurements can be characterised together using the site frequency spectrum (SFS), which plots the number of alleles at each frequency in a population, to quantify the proportion of the genetic richness comprised by alleles at very low frequency (rare alleles) versus those that are common in the population (Allendorf et al. 2012). For example, both rare and common alleles contribute equally to genetic richness, but rare alleles have a negligible impact on heterozygosity (Allendorf 1986). Maintaining both genetic richness and evenness is required to maintain the evolutionary potential of a species (Allendorf et al. 2012), and hence both should be measured to effectively describe the genetic composition of a population (Hoban et al. 2022). Decades of theoretical and empirical studies have demonstrated that genetic diversity is a good predictor of population persistence (Kardos et al. 2021) especially under changing environmental pressures (Sgrò et al. 2011). Thus, when recovering from a bottleneck, populations are at an increased risk of extinction unless their genetic diversity can be restored.

In a closed population, the amount of neutral genetic diversity is determined by the rate of loss due to genetic drift, which is inversely proportional to the effective population size (N_e), and by the rate of gain of novel genetic variation through mutations, which occurs at μ mutations per individual per site per generation. Heterozygosity in a population can be estimated using the mutation-drift evolutionary model (Malécot 1948, Kimura and Crow 1964), a well-established relationship that is described in most textbooks on population and conservation genetics (e.g., Frankham et al. 2010; Hedrick 2011; Allendorf et al. 2012). This model is often used to calculate the equilibrium heterozygosity for neutral sites as $H_{eq} = \frac{4N_e\mu}{4N_e\mu+1}$ (Kimura and Crow 1964). When population sizes change to a new N_e, as is the case for threatened species experiencing population decline or post-decline recovery, genetic diversity starts to move towards the corresponding new equilibrium level. The rate at which this diversity approaches this new equilibrium, is determined by the relative sizes of $2N_e$ and $\frac{1}{2\mu}$. When the two values are sufficiently different, the rate of decay will be effectively controlled by only the smaller component. For example, if a species had a high mutation rate and/or a very large effective population size, such that $2N_e \gg \frac{1}{2\mu}$, the rate of genetic diversity decay and recovery will be inversely proportional to μ (Nei et al. 1975). However, in a species with a low mutation rate and/or a small effective population size, such that $2N_e \ll \frac{1}{2\mu}$, the rate of genetic diversity decay and recovery will be proportional to N_e (Nei et al. 1975).

After a population bottleneck, the recovery of heterozygosity by mutation has been estimated to take millions of generations (Nei et al. 1975; Lande & Barrowclough 1987). Based on these estimates, many population and conservation genetic studies have considered mutation rates too low to substantially contribute to the recovery of genetic diversity in threatened species, at least in the short

term (Lacy 1987; Frankham 2022). However, these seminal papers estimated the recovery of diversity in large populations using rough estimates of the genomic mutation rate, preventing robust estimation of how long heterozygosity would take to recover for threatened species (Nei et al. 1975; Lande & Barrowclough 1987).

Rates of germline single-nucleotide mutations have now been estimated for a diverse range of eukaryotes in parent-offspring and mutation accumulation whole-genome sequencing studies (Wang & Obbard 2023). These include estimates for 116 species, albeit with a skew towards mammals (48 species), with estimates of mutation rates falling between 10^{-10} and 10^{-7} mutations per site per generation (Figure 1a). These precise estimates of mutation rates provide an opportunity to explicitly model the recovery of genetic variation after population bottlenecks in a variety of non-model species across the tree of life. Here we show that, across the known range of mutation rates, the rate of genetic diversity recovery is solely driven by the population size, independent of species-specific mutation rates. Hence, we provide the first universal estimates for the time required to recover genetic diversity after various bottleneck scenarios, in order to contextualise the loss of deep evolutionary histories over short timeframes commensurate with current biodiversity declines. Our results reveal the permanency of genetic diversity losses and highlight the critical need for intervention methods that prevent further loss, or in some cases even increase genetic diversity, to ensure the long-term survival of populations and species.

2 Response of genetic diversity during population-size changes

Using the generalised formula for the inbreeding coefficient (the probability that two randomly selected alleles in a population are identical) while considering genetic drift and mutation (Wright et al. 1931; Malécot 1948; Kimura & Crow 1964) we can derive an equation for heterozygosity as a function of time (generations) as a population moves towards equilibrium (Nei et al. 1975). Heterozygosity, as a proportion (or multiple) of equilibrium heterozygosity ($H_{t/eq}$), can be modelled in terms of time (t, generations), the starting heterozygosity, also as a proportion (or multiple) of the equilibrium heterozygosity ($H_{0/eq}$), the effective population size (N_e) and the mutation rate (μ).

$$H_{t/eq} = 1 - ((1 - 2\mu)\left(1 - \frac{1}{2N_e}\right))^t (1 - H_{0/eq}) \quad (1)$$

When calculating this equation in a population with $N_e = 500$, for a range of mutation rates, we confirm that the time taken to recover heterozygosity in a population is largely unaffected by the mutation rate, except when mutation rate is very high ($> 10^{-4}$ mutations/site/generation; Figure 1b). This means that for the estimated range of germline mutation rates (10^{-10}–10^{-7} mutations/site/generation; Figure 1a), the mutation rate has no impact on the recovery of genetic diversity in a population with $N_e = 500$. For this range of mutation rates, $2N_e \gg \frac{1}{2\mu}$ only occurs when the effective population size is $> 10^6$–10^8, meaning that the mutation rate will only affect the rate of heterozygosity recovery in very large populations. For most species, and particularly those of conservation concern, effective population sizes are smaller by several orders of magnitude; therefore, the rate of genetic diversity recovery, and loss, is not governed by their species-specific mutation rate.

Based on the mutation-drift evolutionary model, we propose a simplified equation for heterozygosity when $2N_e \ll \frac{1}{2\mu}$, which applies to most populations of conservation concern.

$$H_{t/eq} = 1 - \left(1 - \frac{1}{2N_e}\right)^t (1 - H_{0/eq}) \quad (2)$$

The full derivation can be found in the Appendix. Over many generations, $H_{t/eq}$ approaches 1 (H_t approaches H_{eq}) as $\left(1 - \frac{1}{2N_e}\right)^t$ approaches zero. The rate of decay ($1 - \frac{1}{2N_e}$) determines the rate at which the population approaches equilibrium, with the rate being higher for smaller effective

population sizes (N_e). The difference between the starting heterozygosity and the equilibrium heterozygosity ($1 - H_{0/eq}$) also impacts how quickly the heterozygosity approaches equilibrium.

Using this model, we can estimate the recovery of heterozygosity after a bottleneck. A population with N_e = 500, starting with 50% of the equilibrium heterozygosity, takes ~1600 generations to restore diversity to 90% of pre-disturbance level and ~2500 generations to increase diversity to 95% pre-disturbance level (Figure 1b,d). We can also estimate the rate of loss of genetic diversity during a population bottleneck. When starting with a population with N_e = 500 at equilibrium heterozygosity, a population reduction to N_e = 20 will result in the loss of 50% of heterozygosity in only 30 generations, with 80% lost in 80 generations (Figure 1c). With an effective bottleneck size of 5, 50% of the heterozygosity is lost in only 7 generations and 80% in 16 generations. The rate at which genetic diversity approaches equilibrium is determined by the N_e of the population. This explains why genetic diversity can be lost quickly when N_e is reduced during a bottleneck but is recovered slowly after N_e increases. This also explains why more severe bottlenecks with smaller N_e lead to the faster loss of genetic diversity, although they are also approaching a smaller equilibrium value (Figure 1c). These patterns draw attention to the long-term evolutionary consequences of population bottlenecks on heterozygosity.

At mutation–drift equilibrium, genetic richness (e.g., allelic richness) and evenness (e.g., heterozygosity) are correlated, but richness is expected to approach equilibrium more quickly after a change in population size (Nei et al. 1975, Allendorf 1986). Modelling the gain and loss of genetic richness is more challenging than for heterozygosity (Greenbaum et al. 2014), but the effects of bottlenecks on both measures can be observed from genetically explicit forward simulations (Haller & Messer 2023). Results from simulations show that mutation rate does not impact the recovery of heterozygosity nor allelic richness (Figure 2a,b). Our simulations also confirm that allelic richness is generally recovered and lost at much faster rates than heterozygosity (Figure 2c). Following a population decrease, allelic richness is lost faster than heterozygosity for a bottleneck size of N_c = 50 (census population size). In more extreme bottlenecks (N_c < 10), however, heterozygosity can be lost at a similar rate to allelic richness (Supplementary Figure 1). This initial change in richness is driven by a decrease in the number of alleles at very low frequencies (rare alleles), which contribute negligibly to heterozygosity (Figure 2d). Throughout the bottleneck, the SFS shows that the rapid loss of genomic richness after a bottleneck is driven by the disproportionate loss of rare alleles (Figure 2d). More common alleles are lost more slowly, as the SFS returns to its equilibrium distribution. A similar pattern can be observed during the recovery of diversity following a bottleneck. Shortly after the population size returns to N_e = 500, the rapid increase in allelic richness is predominantly driven by an increase in the number of rare alleles. It is not until these new alleles drift to higher frequencies and the SFS returns to equilibrium, over a much longer timeframe, that heterozygosity starts to recover.

Although the recovery times estimated here are too long for conservation management programs to consider, they also represent minimum bounds for the recovery of sufficient diversity for population persistence. Here we have used N_e = 500 in our calculations for the recovery of diversity following a population bottleneck, because it is widely regarded as a benchmark for the minimum effective population size in conservation management (Franklin 1980). However, it has been noted that N_e = 500 does not guarantee sufficient genetic diversity for persistence and that larger effective population sizes (1,000–5,000) are preferred for long-term genetic stability (Frankham et al. 2014). The recovery time for heterozygosity scales directly with N_e, such that recovering the genetic diversity of a population with N_e = 5,000 would take 10 times longer than in a population with N_e = 500. Thus, our estimates represent a lower bound for the recovery time of genetic diversity to healthy levels.

3 The restoration of evolutionary potential

As the recovery of heterozygosity occurs over long timescales, genetically depleted populations will show reduced evolutionary potential in the short term. We have also shown that allelic richness recovers quickly following an expansion in population size, but also that this is unlikely to correspond

to an equivalent increase in evolutionary potential. Standing genetic variation is thought to be more important for adaptation (Barrett & Schluter 2008), because *de novo* mutations start at a low frequency (one copy), while older variants have had time to reach higher frequencies in a population, either through drift or selection. Furthermore, older variants have reduced linkage disequilibrium (are present on more diverse genetic backgrounds) softening selective sweeps (Barrett & Schluter 2008). Standing genetic variation can also increase the likelihood that *de novo* mutations produce new phenotypes, leading to increased evolvability (Tenaillon & Matic 2020). These results suggest that allelic richness is the better indicator of recent demographic processes, and that heterozygosity is the better indicator of the evolutionary potential in a population.

Although we have only demonstrated the timescales for the recovery of neutral diversity, our estimates are likely to provide lower limits for the recovery of non-neutral diversity. The genetic variation that is most likely to contribute to future adaptation are mutations that are neutral or nearly neutral in the current environment and which may provide fitness benefits under future conditions (Barrett & Schluter 2008). Weakly deleterious variants will behave as "effectively neutral" in small populations and hence will recover at rates similar to those of neutral alleles. Some proportion of non-neutral diversity is selected against, so its recovery is expected to be slower than that of neutral diversity (Willi et al. 2006). Therefore, we expect that the recovery of evolutionary potential occurs over a timescale similar to or longer than that of neutral diversity, suggesting that our estimated recovery times are lower bounds.

4 Case study: The demise of the iconic Australian koala

We illustrate the permanency of genetic diversity loss in a threatened species via empirical data for an iconic Australian marsupial, the koala (*Phascolarctos cinereus*). Land clearing across almost the entire east coast of Australia, along with fur trade in the 1900s and ongoing disease epidemics, have driven large reductions in koala population sizes, especially in the southern parts of their range (Adams-Hosking et al. 2016). Recent genomic analyses have found that the Victorian population with $N_e = 19$, contains 58% of the heterozygosity present in the North Queensland population with $N_e = 250$ (McLennan et al. 2024).

We can use the framework presented here to calculate the time that would be needed for the Victorian population to recover to the equilibrium heterozygosity, if the N_e of this population were hypothetically increased to 500. If we assume that the equilibrium value when $N_e = 500$ is at least that seen in the North Queensland population, then we can use $H_{0/eq} = 0.58$ as an upper estimate of the starting heterozygosity in Victoria. We can then use a rearrangement of formula (2) to estimate the time needed for Victoria to regain 90% of the heterozygosity present in a population with $N_e = 500$ (i.e., when $H_{t/eq} = 0.9$). In isolation, the Victorian population would take 1435 generations, or ~ 10,000 years (assuming a generation time of ~7 years) to recover to 90% of this equilibrium genetic diversity from natural mutations alone (full working in Appendix). Our results highlight the need for other measures, such as translocations, to improve the genetic diversity and hence the evolutionary potential of this population.

We have also shown that the easiest way to maintain evolutionary potential is to limit how much is lost during population bottlenecks. Two koala populations in New South Wales have similar effective population sizes to Victoria ($N_e = 27$-28) but have not yet experienced the same level of genetic diversity loss (McLennan et al. 2024). This suggests that population retractions in New South Wales occurred more recently, and recovering N_e in these populations has the potential to prevent the loss of diversity to Victorian levels.

5 Conclusion: so what can we do?

Genetic diversity is being lost by species on a global scale (Shaw et al. 2025); here we have demonstrated that the natural evolutionary process of mutation is insufficient to *restore* genetic

diversity in threatened species. The permanent loss of historical genetic diversity and evolutionary potential is concerning, especially considering the speed at which selective pressures are currently changing. Human-mediated interventions are required to reduce the extinction risk of threatened species and a number of current management practices are able to maximise the maintenance of evolutionary potential with the remaining diversity (Bolam et al. 2023). Recently threatened species, or those with longer generations, may have experienced bottlenecks for a short number of generations, meaning that they have not yet lost a substantial amount of genetic diversity (Nei et al. 1975). Increasing the effective population size for these populations is required to prevent the loss of diversity. Genetically depleted populations can be recovered by increasing the connectivity to, or direct supplementation from, other genetically diverse populations (Clarke et al. 2024) and concerns about outbreeding depression are often overstressed (Chan et al. 2019). Through this method, the effective population size of the entire population can be increased, rather than genetic drift reducing the diversity in each population separately. Assisted gene flow from populations adapted to future climate conditions can further enhance resilience (Aitken & Whitlock 2013).

In cases where there is depleted genetic diversity in the entire species, the only remaining pathways for genetic diversity recovery are more drastic but less proven, and in some cases controversial. Assisted introgression or interspecific hybridisation can introduce valuable genetic diversity to a threatened species and help preserve at least part of the original species' genome (Chan et al. 2019). Finally, introducing historical or novel diversity to the most severely depleted species through synthetic biology may represent a last resort option (Kosch et al. 2022). Our results demonstrate the effective permanency of genetic diversity losses, which highlights the critical need for rapid intervention methods that prevent further loss, or even increase genetic diversity, to ensure the long-term survival of species.

Acknowledgements and Data

The authors have no conflicts of interest. The full derivation of equations presented in this manuscript can be found in the Appendix. Scripts used to simulate population bottlenecks, plot mathematical formulas and summarise previous estimates of mutation rates can be found at github.com/tobykovacs796/diversity_recovery.git.

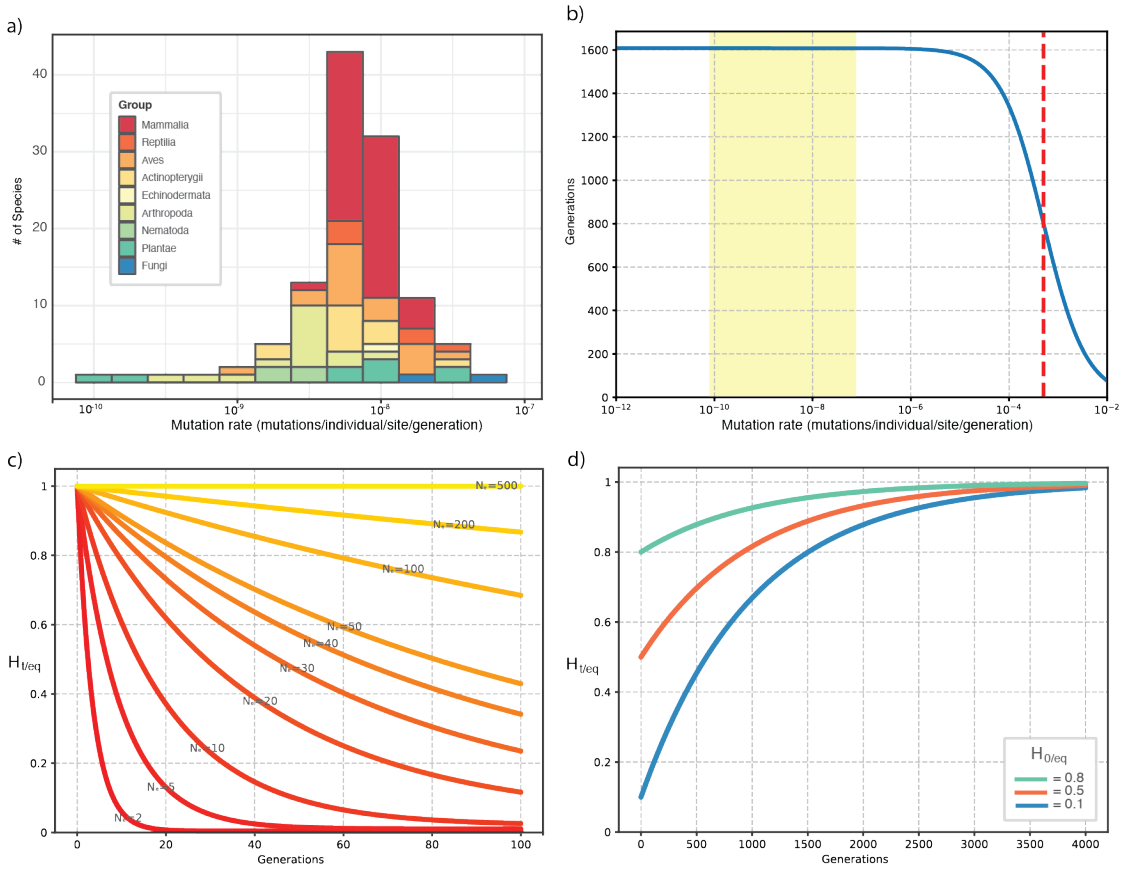


FIGURE 1

(a) Frequency histogram of published mutation rates for multicellular organisms, from the database assembled and updated by Wang and Obbard (2023). **(b)** Number of generations required to recover 90% of the equilibrium heterozygosity in a population with $N_e = 500$, starting with 50% of the equilibrium heterozygosity. This graph uses a rearrangement of formula (1) which calculates heterozygosity as a function of time, the effective population size and the mutation rate. This graph shows the effect that mutation rate has on the time taken to recover heterozygosity, with there being no effect for low mutation rates (when $2N_e \gg \frac{1}{2\mu}$), but a negative relationship for higher mutation rates (when $2N_e \ll \frac{1}{2\mu}$). Yellow shading indicates the range of estimated mutation rates for multicellular organisms, as seen in panel a). The red dashed line indicates $\mu = 0.0005$, which is when $2N_e = \frac{1}{2\mu}$ for $N_e = 500$. Around this value both N_e and μ impact the recovery of genetic variation. **(c)** Loss of heterozygosity during a population bottleneck, calculated using formula (2). The plot shows the loss of genetic diversity in a population with various bottleneck sizes, starting at the equilibrium diversity when $N_e = 500$. The y axis shows heterozygosity divided by the equilibrium heterozygosity for $N_e = 500$. **(d)** Recovery of heterozygosity after a population bottleneck, calculated using formula (2). The plot shows the gain of genetic diversity in a population with $N_e = 500$, starting at 10%, 50% and 80% of the equilibrium heterozygosity (blue, orange and green respectively). The y axis indicates heterozygosity, divided by the equilibrium heterozygosity at $N_e = 500$.

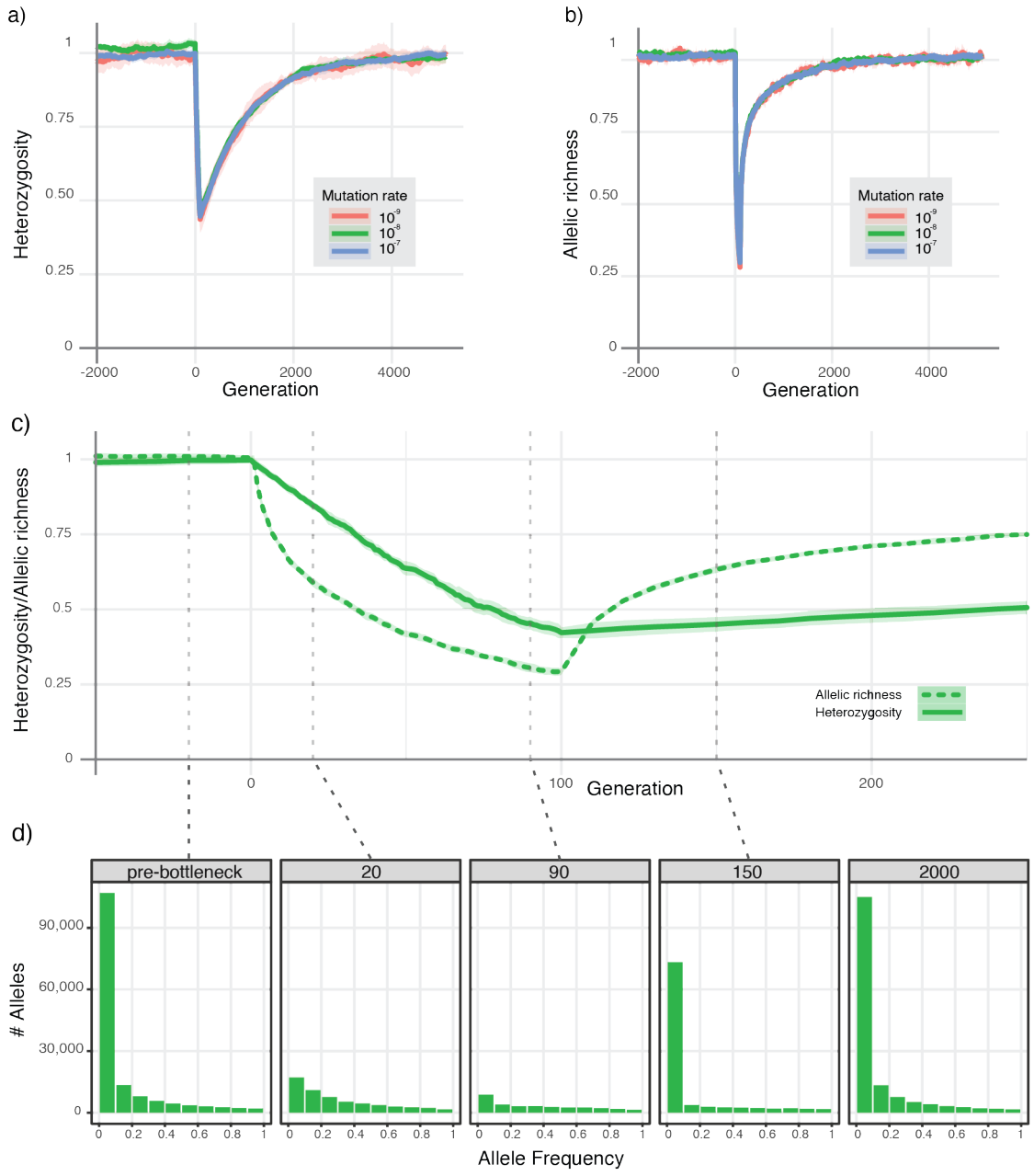
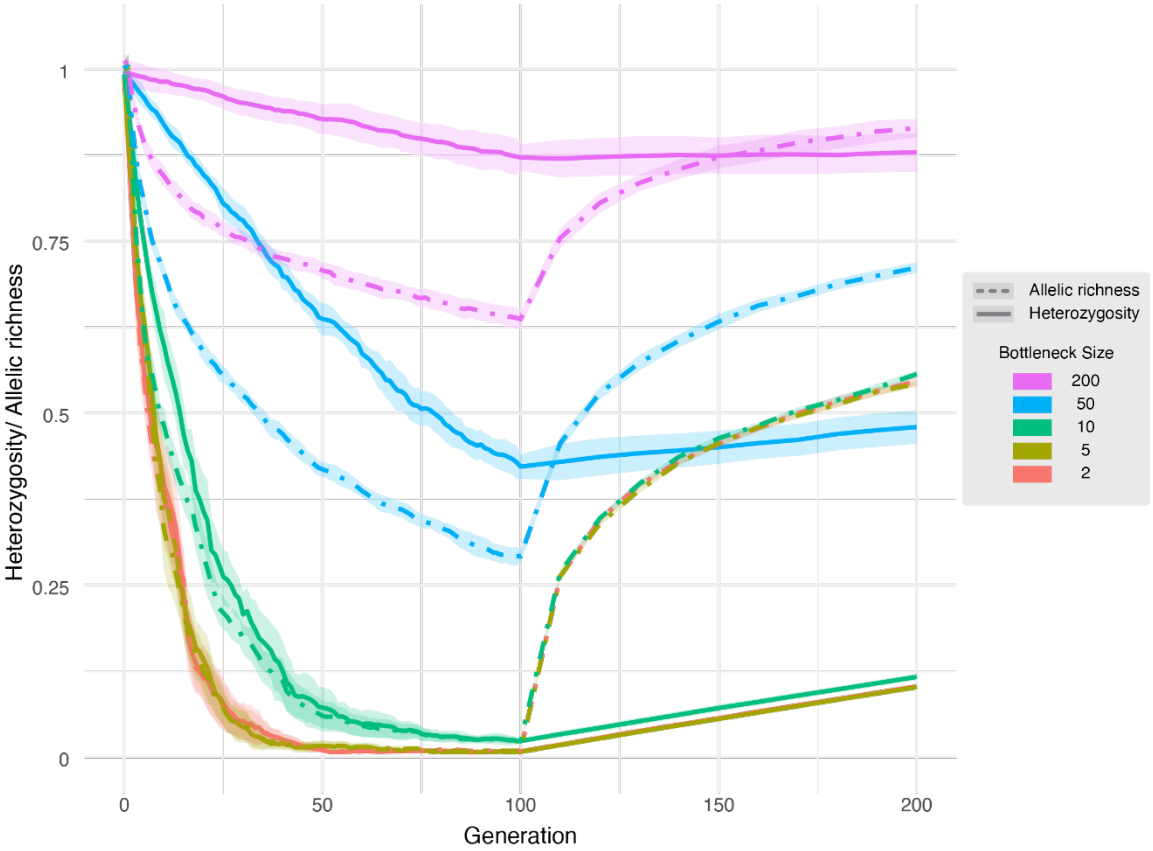


FIGURE 2

Simulations of the loss and recovery of diversity during and after a population bottleneck. Simulations were run with an initial population size of $N_c = 500$ before the population was reduced to $N_c = 50$ (bottleneck) for 100 generations, followed by a return to $N_c = 500$. Ten replicate simulations were run for each scenario. Heterozygosity and allelic richness were averaged across replicates, with pale ribbons indicating 95% confidence intervals. **(a)** Mean heterozygosity across replicates, for three different mutation rates (10^{-7}, 10^{-8}, and 10^{-9} mutations/site/generation) with a bottleneck size of $N_c = 50$. **(b)** Mean allelic richness across replicates, for three different mutation rates (10^{-7}, 10^{-8}, and 10^{-9} mutations/site/generation) with a bottleneck size of $N_c = 50$. **(c)** Comparison of the loss and recovery of allelic richness (dashed line) and heterozygosity (solid line) during and after a bottleneck. Simulations were run with a mutation rate of 10^{-8}. **(d)** Site frequency spectra at five time points through the simulation. Pre-bottleneck, the population is in mutation–drift equilibrium and the SFS has an equilibrium distribution. At 20 and 90 generations after the population-size reduction, there has been a large decrease in allelic richness which is predominantly driven by the loss of low frequency alleles. At 150 generations, shortly after the population-size recovery, there has been a large increase in allelic richness which is predominantly driven by the gain of low-frequency alleles. At 2000

297 generations, drift has had enough time to allow the frequency of rare alleles to increase in the
298 population and the SFS returns to its pre-bottleneck equilibrium distribution.

SUPPLEMENTARY FIGURES



Supplementary Figure 1. *Population simulations of the loss and recovery of allelic richness (dashed line) and heterozygosity (solid line) during and after a population bottleneck. Simulations were run with an initial population size of N_c = 500 before the population size was reduced to N_c = 2 (red), 5 (olive), 10 (green), 50 (blue) and 200 (purple) for 100 generations, followed by a return to N_c = 500. Simulations were run with a mutation rate of 10^{-8} mutations/site/generation.*

References

Adams-Hosking, C., McBride, M.F., Baxter, G., Burgman, M., de Villiers, D., Kavanagh, R., Lawler, I., Lunney, D., Melzer, A., Menkhorst, P., Molsher, R., Moore, B.D., Phalen, D., Rhodes, J.R., Todd, C., Whisson, D. & McAlpine, C.A. (2016). Use of expert knowledge to elicit population trends for the koala (Phascolarctos cinereus). *Divers Distrib*, 22, 249–262.

Aitken, S.N. & Whitlock, M.C. (2013). Assisted gene flow to facilitate local adaptation to climate change. *Annu Rev Ecol Evol Syst*, 44, 367–388.

Allendorf, F.W. (1986). Genetic drift and the loss of alleles versus heterozygosity. *Zoo Biol*, 5, 181–190.

Allendorf, F.W., Aitken, S.N. & Luikart, G.H. (2012). *Conservation and the Genetics of Populations*. Wiley.

Barrett, R.D.H. & Schluter, D. (2008). Adaptation from standing genetic variation. *Trends Ecol Evol*, 23, 38–44.

Bolam, F.C., Ahumada, J., Akçakaya, H.R., Brooks, T.M., Elliott, W., Hoban, S., Mair, L., Mallon, D., McGowan, P.J.K., Raimondo, D., Rodríguez, J.P., Roe, D., Seddon, M.B., Shen, X., Stuart, S.N., Watson, J.E.M. & Butchart, S.H.M. (2023). Over half of threatened species require targeted recovery actions to avert human-induced extinction. *Front Ecol Environ*, 21, 64–70.

CBD. (2022). Decision adopted by the Conference of the Parties to the Convention on Biological Diversity: CBD/COP/DEC/15/4 Kunming- Montreal Global Biodiversity Framework (United Nations Environment Programme, 2022)

Chan, W.Y., Hoffmann, A.A. & van Oppen, M.J.H. (2019). Hybridization as a conservation management tool. *Conserv Lett*, 12, e12652.

Clarke, J.G., Smith, A.C. & Cullingham, C.I. (2024). Genetic rescue often leads to higher fitness as a result of increased heterozygosity across animal taxa. *Mol Ecol*, 33, e17532.

Cowie, R.H., Bouchet, P. & Fontaine, B. (2022). The Sixth Mass Extinction: fact, fiction or speculation? *Biol Rev*, 97, 640–663.

Frankham, R. (2022). Evaluation of proposed genetic goals and targets for the Convention on Biological Diversity. *Conserv Genet*, 23, 865–870.

Frankham, R., Ballou, J.D. & Briscoe, D.A. (2010). *Introduction to Conservation Genetics*. 2nd edn. Cambridge University Press, Cambridge.

Frankham, R., Bradshaw, C.J.A. & Brook, B.W. (2014). Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biol Conserv*, 170, 56–63.

Franklin, I.R. (1980). Evolutionary change in small populations. In: *Conservation biology: an evolutionary-ecological perspective* (eds. Soule, M.E. & Wilcox, B.A.). Sinauer, Sunderland, pp. 135–149.

Greenbaum, G., Templeton, A.R., Zarmi, Y. & Bar-David, S. (2014). Allelic richness following population founding events - A stochastic modeling framework incorporating gene flow and genetic drift. *PLOS One*, 9, e115203.

Haller, B.C. & Messer, P.W. (2023). SLiM 4: Multispecies Eco-Evolutionary Modeling. *Am Nat*, 201, E127–E139.

Hedrick, P.W. (2011). *Genetics of Populations*. 4th edn. Jones & Bartlett Learning, Sudbury, Massachusetts.

Hoban, S., Archer, F.I., Bertola, L.D., Bragg, J.G., Breed, M.F., Bruford, M.W., Coleman, M.A., Ekblom, R., Funk, W.C., Grueber, C.E., Hand, B.K., Jaffé, R., Jensen, E., Johnson, J.S., Kershaw, F., Liggins, L., MacDonald, A.J., Mergeay, J., Miller, J.M., Muller-Karger, F., O'Brien, D., Paz-Vinas, I., Potter, K.M., Razgour, O., Vernesi, C. & Hunter, M.E. (2022). Global genetic diversity status and trends: towards a suite of Essential Biodiversity Variables (EBVs) for genetic composition. *Biol Rev*, 97, 1511–1538.

Kardos, M., Armstrong, E.E., Fitzpatrick, S.W., Hauser, S., Hedrick, P.W., Miller, J.M., Tallmon, D.A. & Funk, C.W. (2021). The crucial role of genome-wide genetic variation in conservation. *PNAS*, 118, e2104642118.

Kimura, M. & Crow, J.F. (1964). The number of neutral alleles maintained in a finite, geographically structured population. *Genetics*, 49, 725–738.

Kosch, T.A., Waddle, A.W., Cooper, C.A., Zenger, K.R., Garrick, D.J., Berger, L. & Skerratt, L.F. (2022). Genetic approaches for increasing fitness in endangered species. *Trends Ecol Evol*, 37, 332–345.

Lacy, R.C. (1987). Loss of Genetic Diversity from Managed Populations: Interacting Effects of Drift, Mutation, Immigration, Selection, and Population Subdivision. *Conserv Biol*, 1, 143–158.

Lande, R. & Barrowclough, G.F. (1987). Effective population size, genetic variation, and their use in population management. In: *Viable Populations for Conservation* (ed. Soule, M.E.). Cambridge University Press, Cambridge, pp. 87–123.

Malécot, G. (1948). *Les mathématiques de l'hérédité*. Paris, France: Masson.

McLennan E.A., Kovacs T.G.L., Silver L.W., Chen Z., Jaya F.R., Ho S.Y.W., Hogg C.J. (2024) Genomics identifies koala populations at risk across eastern Australia. *Ecol Appl*, 35, e3062.

Nei, M., Maruyama, T. & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution*, 29, 1–10.

Scheffers, B.R., De Meester, L., Bridge, T.C.L., Hoffmann, A.A., Pandolfi, J.M., Corlett, R.T., Butchart, S.H.M., Pearce-Kelly, P., Kovacs, K.M., Dudgeon, D., Pacifici, M., Rondinini, C., Foden, W.B., Martin, T.G., Mora, C., Bickford, D. & Watson, J.E.M. (2016). The broad footprint of climate change from genes to biomes to people. *Science*, 354, aaf7671.

Sgrò, C.M., Lowe, A.J. & Hoffmann, A.A. (2011). Building evolutionary resilience for conserving biodiversity under climate change. *Evol Appl*, 4, 326–337.

Shaw, R.E., Farquharson, K.A., Bruford, M.W., Coates, D.J., Elliott, C.P., Mergeay, J., Ottewell, K.M., Segelbacher, G., Hoban, S., Hvilsom, C., Pérez-Espona, S., Ruņģis, D., Aravanopoulos, F., Bertola, L.D., Cotrim, H., Cox, K., Cubric-Curik, V., Ekblom, R., Godoy, J.A., Konopiński, M.K., Laikre, L., Russo, I.-R.M., Veličković, N., Vergeer, P., Vilà, C., Brajkovic, V., Field, D.L., Goodall-Copestake, W.P., Hailer, F., Hopley, T., Zachos, F.E., Alves, P.C., Biedrzycka, A., Binks, R.M., Buiteveld, J., Buzan, E., Byrne, M., Huntley, B., Iacolina, L., Keehnen, N.L.P., Klinga, P., Kopatz, A., Kurland, S., Leonard, J.A., Manfrin, C., Marchesini, A., Millar, M.A., Orozco-terWengel, P., Ottenburghs, J., Posledovich, D., Spencer, P.B., Tourvas, N., Unuk Nahberger, T., van Hooft, P., Verbylaite, R., Vernesi, C. & Grueber, C.E. (2025). Global meta-analysis shows action is needed to halt genetic diversity loss. *Nature*, 638, 704–710.

Tenaillon, O. & Matic, I. (2020). The Impact of Neutral Mutations on Genome Evolvability. *Curr Biol*, 30, R527–R534.

Wang, Y. & Obbard, D.J. (2023). Experimental estimates of germline mutation rate in eukaryotes: a phylogenetic meta-analysis. *Evol Lett*, 7, 216–226.

Willi, Y., Van Buskirk, J. & Hoffmann, A.A. (2006). Limits to the adaptive potential of small populations. *Annu Rev Ecol Evol Syst*, 37, 433–458.

Wright, S. (1931). Evolution in Mendelian Populations. *Genetics*, 16, 97–159.

Appendix to: Genomic data confirms that mutation cannot restore genetic diversity lost through population bottlenecks

Toby G. L. Kovacs, Simon Y. W. Ho, Carolyn J. Hogg, Catherine E. Grueber

June 13, 2025

1 Derivation of H_t, $H_{t/eq}$ and t

1.1 Deriving a formula for Heterozygosity as a function of time as it moves towards equilibrium (H_t)

We want to derive an equation for H_t (the heterozygosity at generation t) in terms of t (generations) for constants μ (mutation rate) and N_e (effective population size). First we will work with F_t (the inbreeding coefficient at generation t) because the equations are simpler, before converting to H_t using the equation: $H_t = 1 - F_t$.

The generalised formula of the inbreeding coefficient based on genetic drift and mutations was derived by Crow and Kimura (1964) as:

$$F_t = \frac{(1-\mu)^2}{2N_e} + \frac{(1-\mu)^2(2N_e-1)}{2N_e}F_{t-1}$$

We recognise that his equation is in the form of the truncated geometric series: $F_t = a + bF_{t-1}$, where:

$$a = \frac{(1-\mu)^2}{2N_e} \text{ and } b = \frac{(1-\mu)^2(2N_e-1)}{2N_e}$$

Based on the geometric series formula, if an equation is in the form $F_t = a + bF_{t-1}$ then we can make a formula for F_t in terms of time (t) and the inbreeding value at $t = 0$ (F_0): $F_t = \frac{a(b^t-1)}{b-1} + b^t F_0$

Substituting into the geometric series formula and then simplifying we get:

$$F_t = \frac{\frac{(1-\mu)^2}{2N_e}\left((\frac{(1-\mu)^2(2N_e-1)}{2N_e})^t - 1\right)}{\frac{(1-\mu)^2(2N_e-1)}{2N_e} - 1} + \left(\frac{(1-\mu)^2(2N_e-1)}{2N_e}\right)^t F_0$$

$$F_t = \frac{\frac{(1-\mu)^2}{2N_e}\left((\frac{(1-\mu)^2(2N_e-1)}{2N_e})^t - 1\right)}{\frac{(1-\mu)^2}{2N_e}\left(2N_e - 1 - \frac{2N_e}{(1-\mu)^2}\right)} + \left(\frac{(1-\mu)^2(2N_e-1)}{2N_e}\right)^t F_0$$

$$F_t = \frac{\left(\frac{(1-\mu)^2(2N_e-1)}{2N_e}\right)^t - 1}{2N_e - 1 - \frac{2N_e}{(1-\mu)^2}} + \left(\frac{(1-\mu)^2(2N_e-1)}{2N_e}\right)^t F_0$$

$$F_t = \frac{\left(\frac{(1-\mu)^2(2N_e-1)}{2N_e}\right)^t}{2N_e - 1 - \frac{2N_e}{(1-\mu)^2}} - \frac{1}{2N_e - 1 - \frac{2N_e}{(1-\mu)^2}} + \left(\frac{(1-\mu)^2(2N_e-1)}{2N_e}\right)^t F_0$$

$$F_t = \left(\frac{(1-\mu)^2(2N_e-1)}{2N_e}\right)^t \left(\frac{1}{2N_e - 1 - \frac{2N_e}{(1-\mu)^2}} + F_0\right) - \frac{1}{2N_e - 1 - \frac{2N_e}{(1-\mu)^2}}$$

We can calculate the Heterozygosity (H_t) from the inbreeding coefficient (F_t) using the equation: $H_t = 1 - F_t$ (and $H_0 = 1 - F_0$) which gives us:

$$H_t = 1 - ((\frac{(1-\mu)^2(2N_e-1)}{2N_e})^t(\frac{1}{2N_e - 1 - \frac{2N_e}{(1-\mu)^2}} + 1 - H_0) - \frac{1}{(2N_e - 1 - \frac{2N_e}{(1-\mu)^2})})$$

which simplifies to

$$H_t = (\frac{(1-\mu)^2(2N_e-1)}{2N_e})^t(H_0 - \frac{1}{2N_e - 1 - \frac{2N_e}{(1-\mu)^2}} - 1) + 1 + \frac{1}{(2N_e - 1 - \frac{2N_e}{(1-\mu)^2})}$$

In the main text we show that for all known metazoa, mutation rates are very small and therefore the square of this number is very very small. We can simplify this equation using approximations assuming that μ is a very small number and $\mu^2 \approx 0$

$$H_t = (\frac{(1-2\mu)(2N_e-1)}{2N_e})^t(H_0 - \frac{1}{2N_e - 1 - \frac{2N_e}{1-2\mu}} - 1) + 1 + \frac{1}{(2N_e - 1 - \frac{2N_e}{1-2\mu})}$$

which simplifies to

$$H_t = ((1-2\mu)(1 - \frac{1}{2N_e}))^t(H_0 - \frac{4N_e}{4N_e + 1 - 2\mu}) + \frac{4N_e}{4N_e + 1 - 2\mu}$$

When considering the gaining of heterozygosity towards the equilibrium, this could be better arranged as:

$$H_t = \frac{4N_e}{4N_e + 1 - 2\mu} - ((1-2\mu)(1 - \frac{1}{2N_e}))^t(\frac{4N_e}{4N_e + 1 - 2\mu} - H_0) \quad (1)$$

This equation demonstrates that there is a exponential decay of the difference between the starting and equilibrium heterozygosity value. The rate of this decay is determined by $(1-2\mu)(1-\frac{1}{2N_e})$ where the impact of N_e and μ depends on their relative sizes. When $\frac{1}{2N_e} \gg 2\mu$ the recovery time is proportional to N_e (the time to recover 90% of equilibrium heterozygosity starting at zero ($H_{0/eq} = 0$) is $\approx 4.6N_e$), but when $\frac{1}{2N_e} \ll 2\mu$ the recovery time is proportional to $\frac{1}{\mu}$ (the time to recover 90% of equilibrium heterozygosity starting at zero ($H_{0/eq} = 0$) is $\approx \frac{1.15}{\mu}$). Note that in the main text, we rearrange these parameters so that both sides of the equation are larger than 1 and easier to think about ($\frac{1}{2\mu} \gg 2N_e$ instead of $\frac{1}{2N_e} \gg 2\mu$).

We can then graph this formula over a range of mutation rates, to see the influence of this parameter (Fig 1b in main text). You can use this equation to test the influence of different values of μ and N_e and plot how neutral genetic diversity changes as a population moves towards equilibrium.

For the known range of genome-wide metazoan mutation rates ($10^{-9} - 10^{-7}$), most populations, and especially those of conservation concern, will have small enough populations such that $\frac{1}{2N_e} \gg 2\mu$ and the time that is required to approach equilibrium heterozygosity is independent of the mutation rate. Hence we can further simplify this equation to:

$$H_t = \frac{4N_e\mu}{4N_e\mu + 1} - (1 - \frac{1}{2N_e})^t(\frac{4N_e\mu}{4N_e\mu + 1} - H_0) \quad (2)$$

1.2 Deriving a formula for the change in Heterozygosity as a proportion of the equilibrium heterozygosity as it moves towards equilibrium ($H_{t/eq}$)

We have shown that the mutation rate does not impact the time it takes for a population to approach equilibrium for most species, however the mutation rate does impact the equilibrium heterozygosity value. It could then be useful to consider heterozygosity in terms of the proportion (or multiple) of the equilibrium heterozygosity, which allows us to predict a universal recovery of heterozygosity across all species. To do this, we convert H_0 and H_t into $H_{0/eq}$ and $H_{t/eq}$ by dividing by the formula for equilibrium heterozygosity given in Crow and Kimura (1964).

$$
\begin{aligned}
H_{0/eq} &= \frac{H_0}{H_{eq}} \\
H_{eq} &= \frac{4N_e\mu}{4N_e\mu+1} \text{ (from Crow and Kimura 1964)} \\
H_{0/eq} &= \frac{H_0}{\left(\frac{4N_e\mu}{4N_e\mu+1}\right)} \\
H_0 &= H_{0/eq}\left(\frac{4N_e\mu}{4N_e\mu+1}\right)
\end{aligned}
$$

We can then substitute this into equation 2

$$
\begin{aligned}
H_t &= \frac{4N_e\mu}{4N_e\mu+1} - \left(1-\frac{1}{2N_e}\right)^t\left(\frac{4N_e\mu}{4N_e\mu+1} - H_{0/eq}\left(\frac{4N_e\mu}{4N_e\mu+1}\right)\right) \\
H_t &= \frac{4N_e\mu}{4N_e\mu+1}\left(1-\left(1-\frac{1}{2N_e}\right)^t(1-H_{0/eq})\right)
\end{aligned}
$$

$$
\begin{aligned}
H_{t/eq} &= \frac{H_t}{H_{eq}} \\
H_{eq} &= \frac{4N_e\mu}{4N_e\mu+1} \quad \text{from Crow and Kimura (1964),} \\
H_{t/eq} &= \frac{\frac{4N_e\mu}{4N_e\mu+1}\left(1-(1-\frac{1}{2N_e})^t(1-H_{0/eq})\right)}{\frac{4N_e\mu}{4N_e\mu+1}}
\end{aligned}
$$

$$
H_{t/eq} = 1-(1-\frac{1}{2N_e})^t(1-H_{0/eq}) \tag{3}
$$

You can use this equation to plot how neutral genetic diversity changes as a population moves towards equilibrium.

We can rearrange this formula to easily calculate the time taken to gain or lose a certain amount of heterozygosity

$$
\begin{aligned}
H_{t/eq} &= 1-(1-\frac{1}{2N_e})^t(1-H_{0/eq}) \\
(1-\frac{1}{2N_e})^t &= \frac{1-H_{t/eq}}{1-H_{0/eq}} \\
\ln\left((1-\frac{1}{2N_e})^t\right) &= \ln\left(|\frac{1-H_{t/eq}}{1-H_{0/eq}}|\right) \\
t\ln\left(1-\frac{1}{2N_e}\right) &= \ln\left(|\frac{1-H_{t/eq}}{1-H_{0/eq}}|\right)
\end{aligned}
$$

$$
t = \frac{\ln\left(|\frac{1-H_{t/eq}}{1-H_{0/eq}}|\right)}{\ln\left(1-\frac{1}{2N_e}\right)} \tag{4}
$$

Calculating how long heterozygosity takes to recover following a population bottleneck: We can use 4 to calculate the time required for the Victorian population of koalas to recover to 90% of the equilibrium diversity ($H_{t/eq} = 0.9$) if the effective population size was increased to 500 ($N_e = 500$). In this example, the Victorian koalas are starting with 58% that estimated in the Northern Queensland population which is likely a conservative lower limit for the equilibrium diversity ($H_{0/eq} = 0.58$).

$$\begin{aligned} t &= \frac{\ln\left(\left|\frac{1-0.9}{1-0.58}\right|\right)}{\ln\left(1-\frac{1}{2\times 500}\right)} \\ &\approx 1435 \text{ generations} \end{aligned}$$

DOUBLECHECKING: To double check that we haven't messed up the algebra, see if after the assumptions, formula 2 can be used to derive the same equilibrium formula as Crow and Kimura (1964).

$$\begin{aligned} &((1-2\mu)(1-\frac{1}{2N_e}))^t \to 0 \text{ because } |(1-2\mu)(1-\frac{1}{2N_e})| < 1 \\ \Rightarrow &\lim_{t\to\infty} H(t) = (0)(H_0 - \frac{4N_e\mu}{4N_e\mu+1}) + \frac{4N_e\mu}{4N_e\mu+1} \\ \Rightarrow &\lim_{t\to\infty} H(t) = \frac{4N_e\mu}{4N_e\mu+1} \end{aligned}$$

This is the same formula derived by Crow and Kimura (1964).

IF WE ONLY USE THE APPROX FOR u2 EQUALS 0: Then we get the following

$$H_t = 1 - \left(\left(\frac{(1-\mu)^2(2N_e-1)}{2N_e}\right)^t \left(\frac{1}{2N_e - 1 - \frac{2N_e}{(1-\mu)^2}} + 1 - H_0\right) - \frac{1}{(2N_e - 1 - \frac{2N_e}{(1-\mu)^2})}\right)$$

$$H_t = 1 - \left(\left((1-2\mu+\mu^2)\left(1-\frac{1}{2N_e}\right)\right)^t \left(\frac{1}{2N_e - 1 - \frac{2N_e}{(1-\mu)^2}} + 1 - H_0\right) - \frac{1}{(2N_e - 1 - \frac{2N_e}{(1-\mu)^2})}\right)$$

$$H_t = 1 - \left(\left((1-2\mu+\mu^2)\left(1-\frac{1}{2N_e}\right)\right)^t \left(\frac{(1-\mu)^2}{(1-\mu)^2(2N_e-1) - 2N_e} + 1 - H_0\right) - \frac{(1-\mu)^2}{(1-\mu)^2(2N_e-1) - 2N_e}\right)$$

$$H_t = 1 - \left(\left((1-2\mu+\mu^2)\left(1-\frac{1}{2N_e}\right)\right)^t \left(\frac{1-2\mu+\mu^2}{(1-2\mu+\mu^2)(2N_e-1) - 2N_e} + 1 - H_0\right) - \frac{1-2\mu+\mu^2}{(1-2\mu+\mu^2)(2N_e-1) - 2N_e}\right)$$

$$\begin{aligned} H_t = &\left((1-2\mu+\mu^2)\left(1-\frac{1}{2N_e}\right)\right)^t \left(H_0 - \frac{1-2\mu+\mu^2+(1-2\mu+\mu^2)(2N_e-1)-2N_e}{(1-2\mu+\mu^2)(2N_e-1)-2N_e}\right) \\ &+ \frac{1-2\mu+\mu^2+(1-2\mu+\mu^2)(2N_e-1)-2N_e}{(1-2\mu+\mu^2)(2N_e-1)-2N_e} \end{aligned}$$

$$H_t = \left((1-2\mu+\mu^2)\left(1-\frac{1}{2N_e}\right)\right)^t \left(H_0 - \frac{2N_e\mu^2-4N_e\mu}{(1-2\mu+\mu^2)(2N_e-1)-2N_e}\right) - \frac{2N_e\mu^2-4N_e\mu}{(1-2\mu+\mu^2)(2N_e-1)-2N_e}$$

$$H_t = \left((1-2\mu+\mu^2)\left(1-\frac{1}{2N_e}\right)\right)^t \left(H_0 - \frac{2N_e\mu^2-4N_e\mu}{2N_e\mu^2-4N_e\mu-1+2\mu-\mu^2}\right) - \frac{2N_e\mu^2-4N_e\mu}{2N_e\mu^2-4N_e\mu-1+2\mu-\mu^2}$$

$$H_t = \left((1-2\mu+\mu^2)\left(1-\frac{1}{2N_e}\right)\right)^t \left(H_0 - \frac{4N_e\mu-2N_e\mu^2}{4N_e\mu+1-2N_e\mu^2-2\mu+\mu^2}\right) - \frac{4N_e\mu-2N_e\mu^2}{4N_e\mu+1-2N_e\mu^2-2\mu+\mu^2}$$