CellPress
OPEN ACCESS

Review

Divergence time shapes gene reuse during repeated adaptation

Magdalena Bohutínská ,[1,2,*,@] and Catherine L. Peichel ,[1]

When diverse lineages repeatedly adapt to similar environmental challenges, the extent to which the same genes are involved (gene reuse) varies across systems. We propose that divergence time among lineages is a key factor driving this variability: as lineages diverge, the extent of gene reuse should decrease due to reductions in allele sharing, functional differentiation among genes, and restructuring of genome architecture. Indeed, we show that many genomic studies of repeated adaptation find that more recently diverged lineages exhibit higher gene reuse during repeated adaptation, but the relationship becomes less clear at older divergence time scales. Thus, future research should explore the factors shaping gene reuse and their interplay across broad divergence time scales for a deeper understanding of evolutionary repeatability.

Can divergence time influence the repeatability of adaptation?

Adaptation is a fundamental process that enables populations and species to cope with environmental challenges. In nature, different **lineages** (see Glossary) sometimes adapt to similar environmental challenges by using the same genes (though not necessarily the same mutations), a phenomenon known as **repeated adaptation** (also referred to as replicated, parallel, or convergent adaptation, as reviewed in [1]) by **gene reuse**. Within this framework, the central focus is on the reuse of genes among different lineages adapting to comparable environments, taking a phenotype-naïve perspective (Boxes 1 and 2). Recently, empirical studies have started to reveal the genomic basis of repeated adaptation in nature, often with the goal of enhancing our ability to predict evolutionary outcomes [2–5]. The rationale is that if a process is repeatable, it should also be predictable (reviewed in [6,7]). Thus, the degree of gene reuse among lineages may be proportional to the predictability of the genetic basis of adaptation.

However, nature is diverse, and there are highly variable levels of gene reuse among lineages [8]. For instance, repeated adaptation to an arctic environment in various species of the Brassicaceae family was associated with a limited amount of gene reuse [9], whereas repeated adaptation to serpentine soil within an *Arabidopsis* species involved a larger number of reused genes [10]. This raises the question of whether the genetic basis of evolutionary processes is too complex or too stochastic to be predictable [7,11], or whether there is an underlying cause for such variability. Four recent studies suggested a promising trend that sheds light on the question of repeatability amidst this variability: among lineages undergoing repeated adaptation, those more recently diverged from each other exhibited a higher degree of gene reuse compared with lineages that diverged longer ago [3,12–14] (Box 3). These case studies complement earlier research that also reported the decreasing repeatability of phenotypic adaptation [15] and of the genetic basis of individual traits [16] with increasing **divergence time**. Indeed, in the earlier examples, the arctic plants showing low gene reuse diverged several million years ago, whereas higher gene reuse in adaptation to serpentine soil was observed within species lineages that

Highlights

Repeated adaptation, also known as parallel or convergent evolution, occurs when different lineages successfully respond to similar environmental challenges.

If the same genes are used by independent lineages during repeated adaptation ('gene reuse'), the genetic basis of adaptation might be predictable.

Recent genomic studies have highlighted that there is variability in the extent of gene reuse among lineages experiencing repeated adaptation.

Divergence time is a promising factor influencing variation in gene reuse in repeated adaptation, as it drives the diversification of shared genetic variation, genome structure, and gene functions among lineages.

Genomic studies of repeated adaptation support the idea that gene reuse decreases with increasing divergence time. However, the relationship is complex, emphasizing the need for additional research to better understand evolutionary repeatability.

[1]Division of Evolutionary Ecology, Institute of Ecology and Evolution, University of Bern, Bern, 3012, Switzerland
[2]Department of Botany, Faculty of Science, Charles University, Prague, 12800, Czech Republic

*Correspondence:
magdalena.bohutinska@natur.cuni.cz
(M. Bohutínská).
@Twitter: @MajdaHolcova

Box 1. Two approaches to study gene reuse in adaptation

Two main approaches are used to study the genetic basis of adaptation: a forward genetic approach that asks which genes underlie a specific adaptive phenotype, and a reverse genetic approach that asks which genes are associated with adaptation to a specific environment. Although the forward genetic approach is specifically focused on phenotypes, the reverse genetic approach is blind to the specific phenotypes that might contribute to adaptation in that environment. These two approaches are complementary [86,89,90] and offer different perspectives on gene reuse.

What can gene reuse during repeated phenotypic evolution tell us?

There are now many studies that have used forward genetic mapping approaches to identify the genes behind repeatedly evolved traits [86,91]. By using this approach, one can ask whether there are particular phenotypes that are more prone to gene reuse and whether there are specific characteristics of the genes that might facilitate its reuse [16,92,93]. For example, it has been proposed that traits governed by only a few genes with a large effect (e.g., pigmentation or xenobiotic resistance [87,94]) will exhibit greater gene reuse during evolution. Conversely, complex traits, characterized by the involvement of numerous genes with a small effect and their interactions (e.g., organ size evolution), may demonstrate lower genetic repeatability [95,96]. Several gene attributes have been linked to an increased likelihood of a gene's reuse in trait evolution, including the high availability of standing genetic variation, low amount of genetic redundancy, location in mutation-prone regions, or functional constraints [16,95,97]. Altogether, understanding the factors that drive gene reuse in repeated phenotypic evolution through forward genetics can contribute to the ongoing discussion concerning the roles of mutational and functional constraints, contingencies in evolution, and the predictability of gene reuse.

What can gene reuse during repeated adaptation tell us?

Many studies have now also used reverse genetic approaches to examine gene reuse among different lineages adapting to similar environmental challenges in nature. In this approach, candidate genes are identified through scans for selection in populations experiencing contrasting environmental conditions [86,98,99]. The identified genes should be those that contribute to all phenotypes involved in adaptation to the specific environment. Thus, genes that contribute to phenotypes with both a simple and complex genetic architecture should be identified. This 'phenotype-naïve' approach is the focus of this review. While here we have reported the impact of divergence time on gene reuse among recently diverged lineages, convincing evidence for other factors influencing gene reuse among repeatedly adapting lineages remains limited. Future research may investigate the roles of ecological strategies (specialist versus generalist lineages) or life history traits (cross- versus self-fertilizing lineages) on the degree of gene reuse (see Outstanding questions). In summary, the focus of reverse genetic studies in repeated adaptation is to understand the factors that cause lineages to adapt in similar ways, informing us about the potentially predictable aspects of lineage evolution. However, they do not provide direct insights into the specific characteristics that promote the reuse of genes or phenotypes in the process of adaptation.

diverged only a few thousand years ago [9,10]. These examples suggest that divergence time might influence the probability of gene reuse in repeated adaptation.

Therefore, here, we first summarize the genetic, phenotypic, and environmental factors that could influence gene reuse across the timescale of divergence. In light of these factors, we propose that the degree of gene reuse in repeated adaptation declines as the lineages diverge over time. Our review of studies examining gene reuse across various divergence times showed that gene reuse indeed decreases with increasing divergence time, particularly for **recent divergences** (< one million generations), whereas the effect remains unclear for **older divergences.**

The mechanisms: factors contributing to divergence-time dependence of gene reuse

A range of genetic, phenotypic, and environmental factors can impact the degree of gene reuse among repeatedly adapting lineages. If these factors are observed to vary with increasing divergence time, they represent candidate mechanisms that may be responsible for shaping the divergence-time dependence of gene reuse. Here, we primarily focus on the genetic factors that potentially contribute to this pattern, specifically the accumulation and maintenance of shared alleles, and the diversification of genome architecture and gene functions.

Glossary

Allele sharing: the phenomenon where two or more populations share the same (adaptive) allele, resulting from recent common ancestry or gene flow.

***De novo* mutations:** mutations that arise independently. These mutations can occur at the same or different genomic positions.

Divergence time: number of generations since two lineages last shared a common ancestor.

Functional diversification: when genes mediate increasingly diverse functions, resulting in the development of distinct phenotypes among lineages. This process can be facilitated by gene duplication.

Gene reuse: involvement of the same genes (but not necessarily the same mutations) in repeated adaptation.

Genome diversification: divergence in mutational and recombinational landscapes, as well as changes in synteny, occurring among lineages as they diverge.

Lineage: a genetically distinguishable unit that can encompass populations, species, or even different kingdoms, representing a branch in the tree of life.

Mutational and recombinational landscapes: the spatial distribution and frequency of mutations and recombination events along the genome.

Older divergence: a divergence time period among lineages where the expectation is that they do not share alleles by descent or introgression. This time frame is roughly equivalent to the divergence time among biological species and beyond. Here, this is arbitrarily defined as more than one million generations.

Orthologous genes, orthologs: genes from different lineages that share a common ancestral gene and can be identified based on their similar DNA or protein sequences.

Recent divergence: a divergence time period among lineages where lineages are genetically distinguishable but may maintain gene flow and share ancestral alleles. Here, this is arbitrarily defined as less than one million generations.

Repeated adaptation: independent improvement of fitness in response to similar environmental pressures in independent lineages. This process can involve the same or different alleles, genes, and/or phenotypes. Also referred to as replicated, parallel, or convergent

Reduction in allele sharing

Gene reuse may be more likely when lineages share beneficial alleles, which can be inherited from a common ancestor as standing genetic variation or introgressed through gene flow (**allele sharing**, Figure 1). The presence of shared adaptive alleles will increase gene reuse because they occur in the same gene, unlike adaptive ***de novo*** mutations, which may occur in the same genes or in different genes with a similar function. Moreover, shared alleles are more likely to be utilized in adaptation than *de novo* mutations, as they are readily available in the adapting lineages at higher frequencies and may have already undergone pretesting for possible deleterious effects in the ancestral population [17]. The probability of allele sharing decreases as lineages diverge due to the accumulation of hybridization barriers and the gradual loss of shared ancestral variation [18–21]. Thus, the degree of gene reuse is predicted to decline with increasing divergence time until the lineages no longer share alleles. At this point, gene reuse is only possible through independent *de novo* mutations in the same gene (Figure 1).

adaptation; note that other authors differentiate between these terms.
Synteny: the preservation of the sequential arrangement of genetic elements, such as genes, in the genomes of distinct lineages.

Box 2. A framework for testing the influence of divergence on gene reuse

To quantify and compare gene reuse in reverse genetic studies of repeated adaptation in light of divergence times among lineages, we propose a framework with three components: study design, candidate gene identification, and testing for significant gene reuse and their relationship to divergence time. Two empirical examples of how a similar framework has been already applied are provided in Box 3 in the main text.

For the study design, a minimum of three naturally occurring lineages (L_{1-3}, Figure I) with varying estimated divergence times ($t_{div,1} > t_{div, 2}$, Figure I) is required. Within each lineage, populations should be adapting to a new (i.e., derived) environment, ideally having colonized the derived environment at a similar point in the past (t_{col}, Figure I). Populations inhabiting the ancestral environment are optimal controls within each lineage.

Next, lineage-specific candidate gene sets that show the signatures of selection in the derived environment are identified (S_{1-3}, Figure I). This should be achieved through a standardized methodology, such as whole-genome sequencing of both the ancestral and derived populations, followed by robust population-level genetic tests to identify the genes with the signatures of selection [98,99].

Finally, the degree of gene reuse (R) between each pair of lineages is estimated by calculating the ratio of the intersection of the candidate gene sets to the union of the candidate gene sets. To determine whether the overlaps between lineages are significant, one can conduct hypergeometric or permutation tests. To evaluate the relationship between t_{div} and R, one can conduct a simple regression analysis, where R serves as the dependent variable and t_{div} as the independent variable. Alternatively, if the case system requires us to account for non-independence among the lineages, one can incorporate a phylogenetic covariance matrix as a random effect; for example, by using phylogenetic generalized least squares (PGLS) regression.
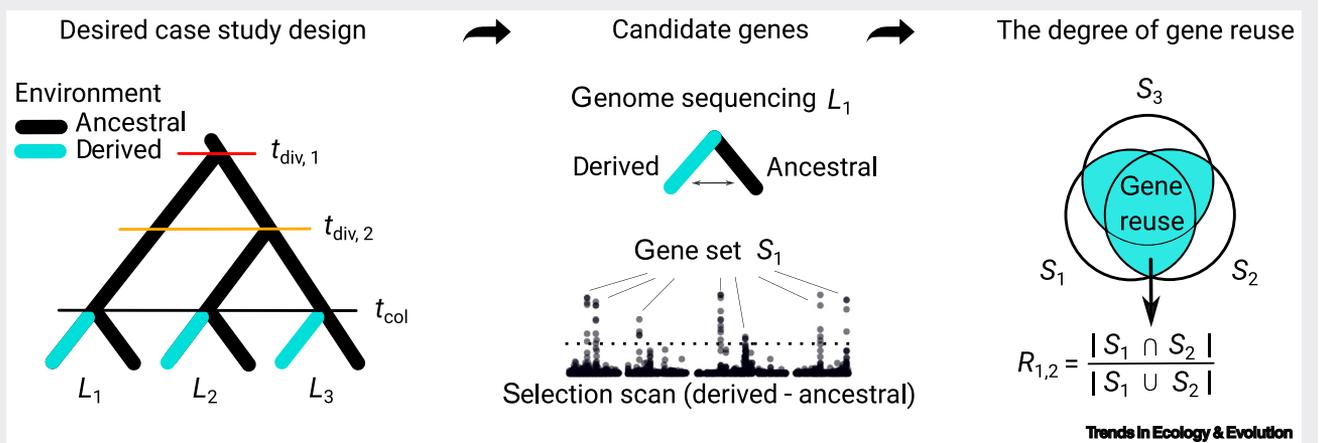

Figure I. Estimating the degree of gene reuse in repeated adaptation to a derived environment. Key: L_{1-3}: three naturally occurring lineages; $t_{div,1}$, $t_{div, 2}$: divergence times; t_{col}: time since the colonization of the derived environment; S_{1-3}: candidate gene sets, showing the signatures of selection in the derived environment within each lineage; $R_{1,2}$: the degree of gene reuse between L_1 and L_2.

Box 3. Exploring gene reuse over a range of divergence times: case studies of repeated adaptation in *Heliconius* and *Arabidopsis*

Two case studies of evolution under similar selection pressure (high elevation) offer valuable insights into the genomic basis of repeated adaptation. The first study focuses on two tropical butterfly species, namely *Heliconius erato* (red postman) and *Heliconius melpomene* (postman butterfly), which have repeatedly adapted to high elevations across South America [14]. By sampling eight lowland and closely related highland populations of both species, the study covers a range of divergence times from recently to 12 million years ago. Genetically and geographically proximal populations exhibit a higher degree of gene reuse in their adaptation to highland environments compared with more distantly related lineages (Figure I). A high level of allele sharing was identified within and between *Heliconius* species, indicating that allele sharing could be a relevant mechanism underlying the observed divergence-time dependency of gene reuse in this system.

The second case study involves two temperate plant species, *Arabidopsis arenosa* (sand rock cress) and *Arabidopsis halleri* (Haller's rock cress), which have repeatedly adapted to high elevations in European mountains [12]. The study analyzed seven highland populations and their related lowland ancestors, showing that gene reuse in highland adaptation decreased with increasing divergence between these lineages, up to the ~600 000 years covered by this system (Figure I). This relationship was mostly explained by a decreasing proportion of repeated selection acting on shared alleles, which were acquired through gene flow or inherited as ancestral standing variation.

Despite a divergence time of 1600 million years between these two systems, we found that around 40% of the genes identified as putatively selected during adaptation to high elevations in *Heliconius* have identifiable orthologs in *Arabidopsis*. However, none of these orthologs was used during adaptation to high altitudes in *Arabidopsis*, indicating a lack of gene reuse, in agreement with the deep divergence time between the two lineages. Altogether, these systems both provide empirical evidence of the importance of divergence time in shaping the degree of gene reuse in repeated adaptation and suggest that decreased allele sharing is the primary underlying factor at these timescales.
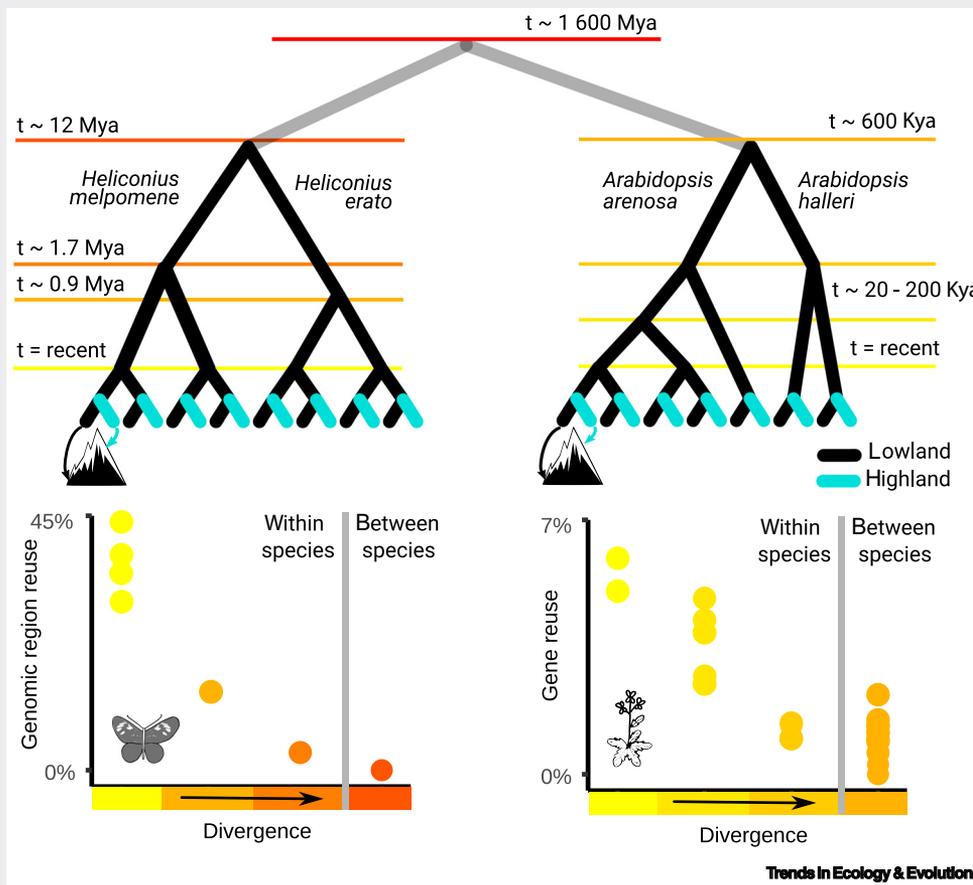

Figure I. Repeated adaptation to high elevation in *Heliconius* butterflies and *Arabidopsis* plants. Highland populations tend to reuse more similar regions when they are less divergent. This trend was observed in both systems despite the differences in their biology and geography. In the *Heliconius* panel, the *y*-axis represents the reuse of genomic regions, as presented in the original article. It shows the percentage of overlap between gene regions initially identified as overlapping among lineages in a lower divergence category. Abbreviations: Kya, thousand years ago; Mya, million years ago.

Genome diversification

Genome architecture, such as **mutational and recombinational landscapes** [22,23] and rearrangement of gene order or **synteny** blocks, also diverges among lineages over time and may play a role in gene reuse (**genome diversification**, Figure 1). For example, changes in
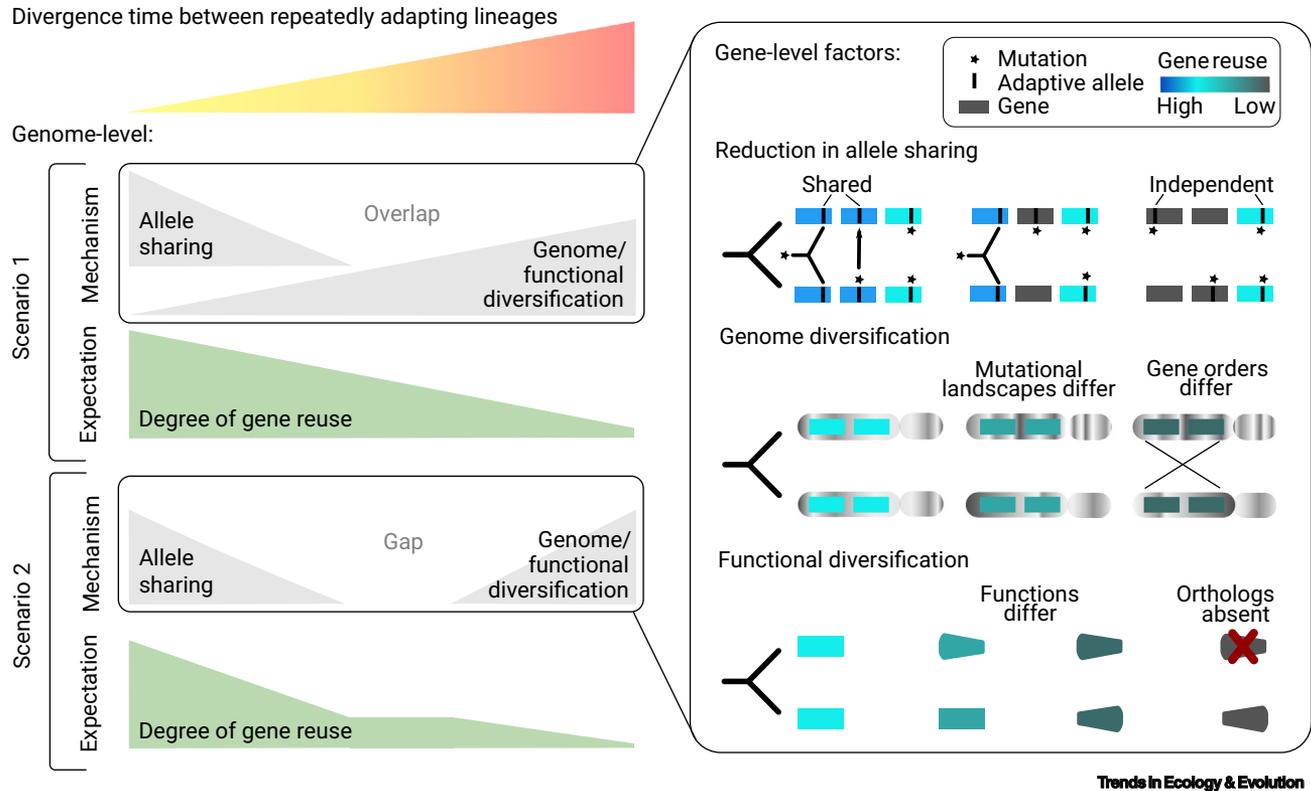
Figure 1. From gene-level factors to genome-wide trends in repeated adaptation across divergence time. On the left side, we present two scenarios of how genetic mechanisms can influence the degree of gene reuse in repeated adaptation with increasing divergence time. Our predictions for genome-wide trends (on the left) are based on the combined genome-wide impacts of gene-level factors (on the right). Note that we depict the presence and direction of the trend but are not making specific predictions about the slope of the relationship at different timepoints. The right side illustrates how reduction in allele sharing and genome/functional diversification influence a decreased probability of reusing individual genes (represented as rectangles with predicted levels of gene reuse, indicated by the color of the rectangle). For each factor, we present genes from two lineages, with their divergence time increasing from left to right.

the mutation or recombination landscape between lineages may lead to a reduction in the likelihood of repeated mutations within the same gene, while changes in gene order may alter the regulatory context of the gene and subsequently lower the probability of gene reuse. Empirical studies have shown that mutation landscapes, encompassing point and structural mutations, as well as the transposition of selfish elements, diversify throughout evolution [24,25]. Furthermore, this diversification was reported to increase with divergence time for point mutations in apes [26]. Similarly, the amount of synteny loss is roughly proportional to divergence time [27]. By contrast, the relationship between divergence time and the diversification of recombination landscapes remains less clear [28,29]. For example, recombination landscapes are conserved among budding yeast species that diverged 15 million years ago [30], whereas there has been a dynamic restructuring of the recombination landscapes between more recently diverging lineages, such as humans and chimpanzees [31]. As a result, this variability can influence the relationship between gene reuse and divergence time.

In conclusion, while our understanding of the evolutionary significance of genomic architecture is still developing, research on the diversification of mutational and synteny landscapes has suggested that they may represent one of the mechanisms leading to decreasing gene reuse with increasing divergence time.

Functional diversification

The diversification of functions among **orthologous genes** (**functional diversification**, Figure 1) also scales with divergence time [32]. As lineages diverge, the functions, biochemical context, gene regulatory networks, and pleiotropic effects of orthologs gradually change, reducing the likelihood of their repeated use for the same adaptive function. Specifically, the similarity in gene expression among tissues and in transcriptional regulation decreases with increasing divergence time [33–35]. The components of gene regulatory networks, as well as their temporal and spatial deployment, also evolve over time [36–39]. Computational estimates and functional work also demonstrated that protein–protein interactions are more conserved within eukaryotic species than between them [40–42]. Finally, orthologs in different species diversify their assigned molecular functions and biological processes over time [33,43,44].

The most extreme case of functional gene diversification occurs when a gene is lost or a new gene emerges in a certain lineage, leading to the absence of a corresponding ortholog in other lineages (Figure 1). For instance, in yeast, only about 30% of genes have orthologs in humans, indicating significant diversification [45]. The gradual loss of orthologs over increasing divergence time is documented in bacteria, archaea, and a set of eukaryotic mitochondrial genes [46,47], suggesting that gene gain and loss is another factor that can contribute to the decreasing degree of gene reuse in repeated adaptation over increasing divergence time.

In summary, as organisms diverge over time, their orthologs diversify in terms of function and presence–absence patterns across lineages. As a result, orthologs can mediate different molecular processes, face distinct genetic constraints, or be replaced by a different set of genes contributing to a particular adaptation. Consequently, it becomes less likely that the same gene would be reused for the same adaptation, which should lead to a gradual reduction in the degree of gene reuse among increasingly divergent lineages.

Phenotypic and ecological factors

Additional factors may influence the degree of gene reuse across a scale of divergence times. For example, lineages may use different phenotypic traits to improve fitness under similar selection pressures. This phenomenon, known as many-to-one mapping of form to function [48], suggests that acquiring a new adaptive function may not necessarily be mediated by the same phenotypic structures; for example, trichomes and darker pigmentation both provide protection against UV radiation at high elevations. As lineages diverge, the phenotypic structures they use for the same adaptation may differ increasingly [15]. Since different structures are likely to have different genetic architectures, including the identity, number, and effect sizes of the genes involved, and different levels of constraints (Box 1), gene reuse becomes less likely with increasing divergence time.

Furthermore, the ecological basis of adaptation may become more diverse with increasing divergence time as a result of a decrease in niche conservatism [49]. More divergent lineages may occupy increasingly dissimilar microhabitats [50]. For example, there may be a significant niche difference between alpine plants and alpine butterflies, even if they are found in a single location, which can reduce the repeatability of selection pressures and render seemingly similar adaptation events no longer comparable. Variable selection pressures may decrease the degree of gene reuse [51], resulting in an even stronger dependence of gene reuse on divergence time [52]. Although decreasing phenotypic and ecological similarity were not the focus of this review, they remain important factors in the relationship between gene reuse and divergence time, as seen in recent studies [3,53].

The expectations: variable interplay of mechanisms across divergence time

The presence of genetic and nongenetic mechanisms that influence gene reuse and scale with divergence time, such as the gradual loss of shared alleles, restructuring of genomic architecture, or diversification of gene functions among diverging lineages, predicts a decrease in gene reuse as lineages diverge. However, not all mechanisms may contribute equally to this pattern across the entire scale of divergence. For instance, allele sharing has its limitations, as lineages where speciation is complete mostly do not share ancestral or introgressed alleles, making it mainly relevant for more recent divergence timescales [21]. Alternatively, significant divergence time may be required for lineages to substantially differentiate gene order or gene functions [54,55], suggesting that functional diversification may predominantly impact longer timescales of divergence.

By taking the timescale and potential interplay among mechanisms into account, we can predict how gene reuse is expected to decline across different segments of the divergence timescale. Figure 1 presents two scenarios to illustrate this concept. In the first scenario, allele sharing overlaps in time with genome and functional diversification mechanisms, leading to a gradual (though not necessarily linear) decrease in gene reuse as divergence time increases. In the second scenario, where the mechanisms do not overlap, the relationship between gene reuse and divergence becomes weak or even nonexistent at intermediate timescales.

To gain deeper insight into the interplay of mechanisms that drive patterns of gene reuse during repeated adaptation, we next examine whether there is empirical evidence for a relationship between gene reuse and divergence across a range of different divergence times and study systems.

The data: divergence time matters

Recent timescales: the effect of divergence through reduction in allele sharing

Our review of case studies on the genomic basis of repeated adaptation provides empirical support for the expectation that there should be a decline in gene reuse as the divergence time increases, when considering recently diverging (< one million generations) lineages (Box 4 and the supplemental information online) [4,5,10,12–14,53,56–64]. Two of these case studies explicitly support our predictions that this decrease in gene reuse is associated with a reduction in allele sharing between lineages (Box 3) [12,14]. This observation is in line with population genetic theory, which
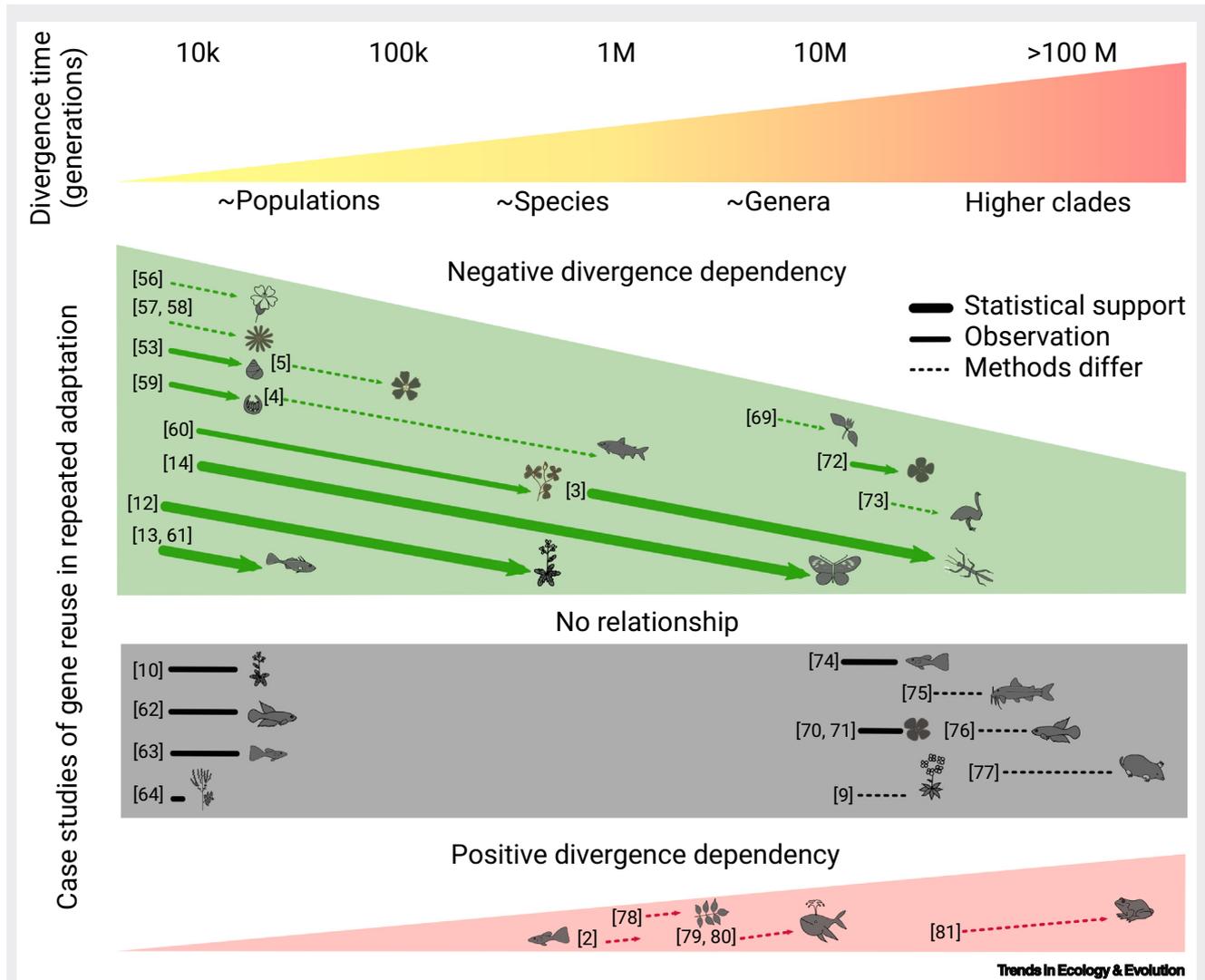

Box 4. The data: divergence matters, especially over recent time scales

To study the relationship between gene reuse and divergence, we searched the literature for genomic studies that examined repeated adaptation across a range of divergence times (following Box 2 in the main text). We focused on genome-wide analyses of adaptation in eukaryotes living in natural environments. In total, we identified 28 case systems (described in 29 published studies and three preprints as of 27 October 2023; note that four of these case systems required the integration of information from two different publications), representing a range of divergence scales spanning from 80 generations to a 100 million generations (see the supplemental information online).

Our analysis revealed that, in most cases, the degree of gene reuse followed the time of divergence between lineages. Specifically, 18 out of 28 case studies showed a relationship between divergence time and gene reuse (Figure I). In 14 studies, there was less gene reuse at greater divergence times (negative divergence dependency), while in four studies, there was higher gene reuse at greater divergence times (positive divergence dependency) (significant enrichment of negative relationships, $\chi 2 = 8.4$, df = 3, $P = 0.04$). Furthermore, all four studies that explicitly focused on this relationship showed negative divergence dependency [3,12–14].

However, the relationship varied depending on the range of divergence times covered by each study. At the short end of the time scale, covering divergences between populations and congeneric species (divergence less than one million generations), 10 out of 14 available studies supported a negative divergence dependency. The remaining four studies showed an ambiguous pattern and none of them showed positive divergence dependency (significant enrichment of negative relationships, $\chi 2 = 16.4$, df = 3, $P < 0.001$; Figure I). This suggests that divergence is a significant factor driving the degree of gene reuse at shallow levels of divergence.

When comparing lineages that diverged long ago, we found no evidence that gene reuse is dependent on divergence time. In 14 studies of lineages diverging more than one million generations ago, we observed a range of results: four studies showed positive divergence dependency, four showed negative dependency, and the remaining six showed ambiguous trends ($\chi 2 = 1.6$, df = 3, $P < 0.65$; Figure I). Overall, we did not find support for divergence-time dependency of gene reuse at older divergence scales.

Figure I. The relationship between divergence time and the degree of gene reuse, and case studies that show a negative relationship (green), a positive relationship (red), or no clear relationship (black). The arrows indicate the direction of the relationship, with divergence shown on a logarithmic scale, indicated as the approximate number of generations separating the studied lineages. Bold lines indicate a significant relationship between gene reuse and divergence estimated by the corresponding case study (indicated by its reference number). Broken lines highlight studies which differ in a certain aspect from our desired setup (see Box 2 in the main text), most commonly because they used reduced representation sequencing or lacked the closely related ancestral lineage. See the supplemental information online for details about the case studies and how they were selected. See [2–5,9,10,12–14,53,56–64,69–81].

proposes that shared alleles are lost within the number of generations roughly corresponding to nine times the effective population size [65]. Considering the effective population sizes reported in our dataset, which range from thousands to hundreds of thousands of individuals [13,63], we would predict the loss of most shared alleles to occur within the first few million generations of separation, consistent with the divergence-time dependency observed in these studies. Additionally, historical bottlenecks, higher mutation rates, or low levels of balancing selection can further accelerate this process [66–68], which may provide an explanation for the ambiguous patterns observed in some of the case studies with similar divergence times (Box 4 and supplemental information)

[10,62–64]. In summary, a decrease in allele sharing, which necessitates independent mutations for adaptation, emerges as a significant genetic mechanism contributing to the negative relationship between divergence time and gene reuse among recently diverged lineages.

Old timescales: the unclear effect of divergence despite genome and functional diversification

By contrast, our analysis of longer-diverged lineages (between 1 and 100 million generations) revealed a complex pattern in the relationship between gene reuse and divergence time (Box 4 and supplemental information) [2,3,9,69–81]. This complexity has arisen partly from differences in study designs across older divergence times compared with our proposed 'optimal' design in Box 2. For example, studies of the repeated adaptation of marine mammals to an underwater lifestyle lack a closely related 'terrestrial whale' population, leading researchers to use the closest available lineages (i.e., cows). This presents challenges in identifying positively selected genes and quantifying gene reuse, potentially resulting in ambiguous findings regarding the relationship between gene reuse and divergence time [82].

Alternatively, many of these case studies of repeated adaptation may span a timeframe with a lack of shared alleles but with negligible genome/functional diversification. This could result in limited divergence time-dependency of gene reuse at some divergence levels, as depicted in Scenario 2 in Figure 1. Several studies providing the timing of genome/functional diversification may support this scenario. First, genome diversification is likely initiated early during divergence of a lineage, but significant differences may take time to accumulate. Simulations and empirical work have suggested the translocation of a few genes per million years [83] and the gradual loss of ancestral gene order over several million years [27]. For example, there is still 99% conservation of synteny in *Drosophila* species that diverged 35 million years ago, but only approximately 10% conservation between flies and honeybees, which diverged 350 million years ago [84]. Second, studies suggest early onset of functional diversification for some, but not for all genes. Simulations suggest early functional diversification with one to ten changes in protein–protein interactions every 10 000 years of divergence within eukaryotic species [40,41], and functional work has shown that 7% of genes have functionally diversified in *Caenorhabditis* species that diverged 40 million years ago [85]. However, essential genes may remain functionally unchanged for a longer time. For instance, half of essential yeast genes are functionally interchangeable with their human orthologs after over 1 billion years of divergence [45].

Thus, when evaluating the impact and timing of functional diversification on gene reuse in repeated adaptation, we need to consider the extent to which adaptation relies on conserved molecular pathways. If adaptation heavily depends on conserved pathways, the initial impact of functional diversification may be negligible, contributing to the gap between allele sharing and functional diversification (Scenario 2 in Figure 1). This could be the driver of the ambiguous relationship between gene reuse and divergence across the older timescales observed in our case studies (Box 4). However, the interaction between genome and functional diversification, and reduction in allele sharing are also likely to vary depending on the species. Larger effective population sizes may result in a slower decay of shared alleles, supporting Scenario 1 (Figure 1). Alternatively, species with smaller or fluctuating effective population sizes or those relying on conserved molecular pathways for adaptation may lose shared alleles before their gene functions diverge, supporting Scenario 2 (Figure 1). Currently, the divergence time scope of our review does not enable us to differentiate between Scenarios 1 and 2, as the maximum difference in divergence time is 54 million generations (Box 4 and case study in [81]), and we have limited information about the timing of underlying mechanisms. However, by conducting studies that cover a broader range of divergence scales, we may gain further insights into the relationship between gene reuse and divergence, as well as the underlying mechanisms involved (see Outstanding questions).

Concluding remarks

Recent genomic studies have significantly advanced our understanding of the genomic basis of repeated adaptation. In this review, we hypothesized that gene reuse would decrease as lineages diverge, in light of factors such as reduced allele sharing and the diversification of gene function. However, a comprehensive analysis of available genomic studies emphasized the impact of divergence time on gene reuse across recent, but not older, timescales. Thus, our current understanding of how the degree of gene reuse varies across a broader range of divergences remains limited.

To expand our knowledge, research on the genomic basis of repeated adaptation in nature should include a wider range of divergence times, integrating population genetics and large-scale macroevolutionary processes. Intermediate levels of divergence, where the mechanisms of allele sharing and genome/functional diversification may operate together, are particularly intriguing. However, finding selection pressures that act on the same genes over long enough timescales might be difficult. Studying adaptation to challenges such as polyploidy, temperature, or pathogens, which might directly target conserved cellular processes [78,86], or adaptations that require modifications of highly conserved molecular machineries, such as pigment production pathways [87] or ion transporters [88], could serve as promising model systems.

Furthermore, the direct relationship between genome and functional diversification and the extent of gene reuse across varying levels of divergence remains to be established. Further, there is a need to explore additional genetic and nongenetic mechanisms influencing the relationship between gene reuse and divergence time. Determining these factors and their timing will provide valuable insights into the drivers of evolutionary repeatability.

Acknowledgments

Declaration of interests

The authors have no interests to declare.

Supplemental information

Supplemental information associated with this article can be found online https://doi.org/10.1016/j.tree.2023.11.007.

References

1. Cerca, J. (2023) Understanding natural selection and similarity: convergent, parallel, and repeated evolution. *Mol. Ecol.* 32, 5451–5462
2. Brown, A.P. *et al.* (2019) Local ancestry analysis reveals genomic convergence in extremophile fishes. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 374, 20180240
3. Chaturvedi, S. *et al.* (2022) Climatic similarity and genomic background shape the extent of parallel adaptation in *Timema* stick insects. *Nat. Ecol. Evol.* 6, 1952–1964
4. Jacobs, A. *et al.* (2020) Parallelism in eco-morphology and gene expression despite variable evolutionary and genomic backgrounds in a Holarctic fish. *PLoS Genet.* 16, e1008658
5. Szukala, A. *et al.* (2022) Polygenic routes lead to parallel altitudinal adaptation in *Heliosperma pusillum* (Caryophyllaceae). *Mol. Ecol.* 32, 1832–1847
6. Blount, Z.D. *et al.* (2018) Contingency and determinism in evolution: replaying life's tape. *Science* 362, eaam5979
7. Gompert, Z. *et al.* (2022) Laplace's demon in biology: models of evolutionary prediction. *Evolution* 76, 2794–2810
8. Waters, J.M. and McCulloch, G.A. (2021) Reinventing the wheel? Reassessing the roles of gene flow, sorting and convergence in repeated evolution. *Mol. Ecol.* 30, 4162–4172
9. Birkeland, S. *et al.* (2020) Multiple genetic trajectories to extreme abiotic stress adaptation in arctic Brassicaceae. *Mol. Biol. Evol.* 37, 2052–2068
10. Konečná, V. *et al.* (2021) Parallel adaptation in autopolyploid *Arabidopsis arenosa* is dominated by repeated recruitment of shared alleles. *Nat. Commun.* 12, 4979
11. Lewontin, R.C. (1966) Is nature probable or capricious? *BioScience* 16, 25–27
12. Bohutínská, M. *et al.* (2021) Genomic basis of parallel adaptation varies with divergence in Arabidopsis and its relatives. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2022713118
13. Magalhaes, I.S. *et al.* (2021) Intercontinental genomic parallelism in multiple three-spined stickleback adaptive radiations. *Nat. Ecol. Evol.* 5, 251–261
14. Montejo-Kovacevich, G. *et al.* (2022) Repeated genetic adaptation to altitude in two tropical butterflies. *Nat. Commun.* 13, 4676

15. Ord, T.J. and Summers, T.C. (2015) Repeated evolution and the impact of evolutionary history on adaptation. *BMC Evol. Biol.* 15, 137

16. Conte, G.L. *et al.* (2012) The probability of genetic parallelism and convergence in natural populations. *Proc. R. Soc. B* 279, 5039–5047

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

18. Edwards, S. and Beerli, P. (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* 54, 1839–1854

19. Mallet, J. *et al.* (2016) How reticulated are species? *BioEssays* 38, 140–149

20. Mendes, F. *et al.* (2016) Gene tree discordance can generate patterns of diminishing convergence over time. *Mol. Biol. Evol.* 33, 3299–3307

21. Roux, C. *et al.* (2016) Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLoS Biol.* 14, e2000234

22. Peñalba, J.V. and Wolf, J.B.W. (2020) From molecules to populations: appreciating and estimating recombination rate variation. *Nat. Rev. Genet.* 21, 476–492

23. Sasaki, A. and Nowak, M.A. (2003) Mutation landscapes. *J. Theor. Biol.* 224, 241–247

24. Gao, Z. *et al.* (2023) Limited role of generation time changes in driving the evolution of the mutation spectrum in humans. *eLife* 12, e81188

25. Schrader, L. and Schmitz, J. (2019) The impact of transposable elements in adaptive evolution. *Mol. Ecol.* 28, 1537–1549

26. Goldberg, M.E. and Harris, K. (2021) Mutational signatures of replication timing and epigenetic modification persist through the global divergence of mutation spectra across the great ape phylogeny. *Genome Biol. Evol.* 14, evab104

27. Koonin, E.V. (2009) Evolution of genome architecture. *Int. J. Biochem. Cell Biol.* 41, 298–306

28. Brazier, T. and Glémin, S. (2022) Diversity and determinants of recombination landscapes in flowering plants. *PLoS Genet.* 18, e1010141

29. Smukowski, C.S. and Noor, M.A.F. (2011) Recombination rate variation in closely related species. *Heredity* 107, 496–508

30. Lam, I. and Keeney, S. (2015) Nonparadoxical evolutionary stability of the recombination initiation landscape in yeast. *Science* 350, 932–937

31. Dapper, A.L. and Payseur, B.A. (2017) Connecting theory and data to understand recombination rate evolution. *Philos. Trans. R. Soc. B Biol. Sci.* 372, 20160469

32. Gabaldón, T. and Koonin, E.V. (2013) Functional and evolutionary implications of gene orthology. *Nat. Rev. Genet.* 14, 360–366

33. Chen, X. and Zhang, J. (2012) The ortholog conjecture is untestable by the current gene ontology but is supported by RNA sequencing data. *PLoS Comput. Biol.* 8, e1002784

34. Martin, T. and Fraser, H.B. (2018) Comparative expression profiling reveals widespread coordinated evolution of gene expression across eukaryotes. *Nat. Commun.* 9, 4963

35. Thompson, D.A. *et al.* (2013) Evolutionary principles of modular gene regulation in yeasts. *eLife* 2, e00603

36. Rogers, J.M. and Bulyk, M.L. (2018) Diversification of transcription factor–DNA interactions and the evolution of gene regulatory networks. *Wiley Interdiscip. Rev. Syst. Biol. Med.* 10, e1423

37. Gaudinier, A. and Blackman, B.K. (2020) Evolutionary processes from the perspective of flowering time diversity. *New Phytol.* 225, 1883–1898

38. McQueen, E. and Rebeiz, M. (2020) On the individuality of gene regulatory networks: how does network re-use affect subsequent evolution? *Curr. Top. Dev. Biol.* 139, 375–405

39. Hill, M.S. *et al.* (2021) Molecular and evolutionary processes governing variation in gene expression. *Nat. Rev. Genet.* 22, 203–215

40. Beltrao, P. and Serrano, L. (2007) Specificity and evolvability in eukaryotic protein interaction networks. *PLoS Comput. Biol.* 3, e25

41. Mika, S. and Rost, B. (2006) Protein–protein interactions more conserved within species than across species. *PLoS Comput. Biol.* 2, e79

42. Ryan, C.J. *et al.* (2012) Hierarchical modularity and the evolution of genetic interactomes across species. *Mol. Cell* 46, 691–704

43. Altenhoff, A.M. *et al.* (2021) OMA orthology in 2021: website overhaul, conserved isoforms, ancestral gene order and more. *Nucleic Acids Res.* 49, D373–D379

44. Nehrt, N.L. *et al.* (2011) Testing the ortholog conjecture with comparative functional genomic data from mammals. *PLoS Comput. Biol.* 7, e1002073

45. Kachroo, A.H. *et al.* (2015) Systematic humanization of yeast genes reveals conserved functions and genetic modularity. *Science* 348, 921–925

46. Li, Y. *et al.* (2014) Expansion of biological pathways based on evolutionary inference. *Cell* 158, 213–225

47. Wolf, Y.I. and Koonin, E.V. (2012) A tight link between orthologs and bidirectional best hits in bacterial and archaeal genomes. *Genome Biol. Evol.* 4, 1286–1294

48. Wainwright, P.C. *et al.* (2005) Many-to-one mapping of form to function: a general principle in organismal design? *Int. Comp. Biol.* 45, 256–262

49. Losos, J.B. (2008) Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecol. Lett.* 11, 995–1003

50. Crisp, M.D. and Cook, L.G. (2012) Phylogenetic niche conservatism: what are the underlying evolutionary and ecological causes? *New Phytol.* 196, 681–694

51. Stuart, Y. *et al.* (2017) Contrasting effects of environment and genetics generate a continuum of parallel evolution. *Nat. Ecol. Evol.* 1, 0158

52. Thompson, K.A. *et al.* (2019) Parallel genetic evolution and speciation from standing variation. *Evol. Lett.* 3, 129–141

53. Morales, H.E. *et al.* (2019) Genomic architecture of parallel ecological divergence: beyond a single environmental contrast. *Sci. Adv.* 5, eaav9963

54. Lai, H. *et al.* (2023) Multiple intermolecular interactions facilitate rapid evolution of essential genes. *Nat. Ecol. Evol.* 7, 745–755

55. Simakov, O. *et al.* (2022) Deeply conserved synteny and the evolution of metazoan chromosomes. *Sci. Adv.* 8, eabi5884

56. Papadopulos, A.S.T. *et al.* (2021) Rapid parallel adaptation to anthropogenic heavy metal pollution. *Mol. Biol. Evol.* 38, 3724–3736

57. Roda, F. *et al.* (2013) Convergence and divergence during the adaptation to similar environments by an Australian groundsel. *Evolution* 67, 2515–2529

58. James, M.E. *et al.* (2021) Highly replicated evolution of parapatric ecotypes. *Mol. Biol. Evol.* 38, 4805–4821

59. Hartmann, F.E. *et al.* (2021) The complex genomic basis of rapid convergent adaptation to pesticides across continents in a fungal plant pathogen. *Mol. Ecol.* 30, 5390–5405

60. Preite, V. *et al.* (2019) Convergent evolution in *Arabidopsis halleri* and *Arabidopsis arenosa* on calamine metalliferous soils. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 374, 20180243

61. Fang, B. *et al.* (2020) On the causes of geographically heterogenenous parallel evolution in sticklebacks. *Nat. Ecol. Evol.* 4, 1105–1115

62. Reid, N.M. *et al.* (2016) The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science* 354, 1305–1308

63. Whiting, J.R. *et al.* (2021) Drainage-structuring of ancestral variation and a common functional pathway shape limited genomic convergence in natural high- and low-predation guppies. *PLoS Genet.* 17, e1009566

64. Van Boheemen, L.A. and Hodgins, K.A. (2020) Rapid repeatable phenotypic and genomic adaptation following multiple introductions. *Mol. Ecol.* 29, 4102–4117

65. Hudson, R.R. and Coyne, J.A. (2002) Mathematical consequences of the genealogical species concept. *Evolution* 56, 1557–1565

66. Guerrero, R.F. and Hahn, M.W. (2017) Speciation as a sieve for ancestral polymorphism. *Mol. Ecol.* 26, 5362–5368

67. Maddison, W.P. (1997) Gene trees in species trees. *Syst. Biol.* 46, 523–536

68. Stern, D.B. and Lee, C.E. (2020) Evolutionary origins of genomic adaptations in an invasive copepod. *Nat. Ecol. Evol.* 4, 1084–1094

69. Xu, S. *et al.* (2017) Genome-wide convergence during evolution of mangroves from woody plants. *Mol. Biol. Evol.* 34, 1008–1015

70. Bray, S.M. *et al.* (2023) Kinetochore and ionomic adaptation to whole genome duplication. *BioRxiv* Published online September 28, 2023. https://doi.org/10.1101/2023.09.27.559727

71. Bohutínská, M. *et al.* (2021) Novelty and convergence in adaptation to whole genome duplication. *Mol. Biol. Evol.* 38, 3910–3924

72. Rellstab, C. *et al.* (2020) Genomic signatures of convergent adaptation to alpine environments in three Brassicaceae species. *Mol. Ecol.* 29, 4350–4365

73. Sackton, T.B. *et al.* (2019) Convergent regulatory evolution and loss of flight in paleognathous birds. *Science* 364, 74–78

74. Greenway, R. *et al.* (2021) Convergent adaptation and ecological speciation result from unique genomic mechanisms in sympatric extremophile fishes. *BioRxiv* Published online June 28, 2021. https://doi.org/10.1101/2021.06.28.450104

75. Yang, L. *et al.* (2021) Genomic and functional evidence reveals convergent evolution in fishes on the Tibetan Plateau. *Mol. Ecol.* 30, 5752–5764

76. Thompson, A.W. *et al.* (2021) Deterministic shifts in molecular evolution correlate with convergence to annualism in killifishes. *BioRxiv* Published online August 11, 2021. https://doi.org/10.1101/2021.08.09.455723

77. Davies, K.T.J. *et al.* (2018) Limited evidence for parallel molecular adaptations associated with the subterranean niche in mammals: a comparative study of three superorders. *Mol. Biol. Evol.* 35, 2544–2559

78. Kelly, L.J. *et al.* (2020) Convergent molecular evolution among ash species resistant to the emerald ash borer. *Nat. Ecol. Evol.* 4, 1116–1128

79. Foote, A.D. *et al.* (2015) Convergent evolution of the genomes of marine mammals. *Nat. Genet.* 47, 272–275

80. Yuan, Y. *et al.* (2021) Comparative genomics provides insights into the aquatic adaptations of mammals. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2106080118

81. Lu, B. *et al.* (2020) Molecular convergent and parallel evolution among four high-elevation anuran species from the Tibetan region. *BMC Genomics* 21, 839

82. Velotta, J.P. *et al.* (2022) Repeated genetic targets of natural selection underlying adaptation of fishes to changing salinity. *Int. Comp. Biol.* 62, 357–375

83. Lv, J. *et al.* (2011) Constraints on genes shape long-term conservation of macro-synteny in metazoan genomes. *BMC Bioinformatics* 12, S11

84. Zdobnov, E.M. and Bork, P. (2007) Quantification of insect genome divergence. *Trends Genet.* 23, 16–20

85. Verster, A.J. *et al.* (2014) Comparative RNAi screens in *C. elegans* and *C. briggsae* reveal the impact of developmental system drift on gene function. *PLoS Genet.* 10, e1004077

86. Bomblies, K. and Peichel, C.L. (2022) Genetics of adaptation. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2122152119

87. Elkin, J. *et al.* (2023) Analysis of the genetic loci of pigment pattern evolution in vertebrates. *Biol. Rev. Camb. Philos. Soc.* 98, 1250–1277

88. Konečná, V. *et al.* (2020) The evolutionary genomics of serpentine adaptation. *Front. Plant Sci.* 11, 574616

89. Stinchcombe, J.R. and Hoekstra, H.E. (2008) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity* 100, 158–170

90. Barrett, R.D.H. and Hoekstra, H.E. (2011) Molecular spandrels: tests of adaptation at the genetic level. *Nat. Rev. Genet.* 12 767-180

91. Courtier-Orgogozo, V. *et al.* (2020) Gephebase, a database of genotype-phenotype relationships for natural and domesticated variation in Eukaryotes. *Nucleic Acids Res.* 48, D696–D703

92. Stern, D. (2010) *Evolution, Development, and the Predictable Genome*, Roberts and Company

93. Martin, A. and Orgogozo, V. (2013) The loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution* 67, 1235–1250

94. Feldman, C.R. *et al.* (2012) Constraint shapes convergence in tetrodotoxin-resistant sodium channels of snakes. *Proc. Natl. Acad. Sci. U. S. A.* 109, 4556–4561

95. Láruson, A.J. *et al.* (2020) The importance of genetic redundancy in evolution. *Trends Ecol. Evol.* 35, 809–822

96. James, M.E. *et al.* (2023) Replicated evolution in plants. *Annu. Rev. Plant Biol.* 74, 697–725

97. Yeaman, S. *et al.* (2018) Quantifying how constraints limit the diversity of viable routes to adaptation. *PLoS Genet.* 14, e1007717

98. Hoban, S. *et al.* (2016) Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *Am. Nat.* 188, 379–397

99. Tanaka, T. *et al.* (2023) Power of neutrality tests for detecting natural selection. *G3* 13, jkad161