PAPER

From Valleys to Peaks: The Role of Evolvability in Fitness Landscape Navigation

Priyanka Mehra[a,] and Arend Hintze Author[a, b, *]

[a]Department of MicroData Analysis, Dalarna University, Högskolegatan 2, 791 88, Falun, Sweden and [b]BEACON Center for the Study of Evolution in Action, Michigan State University, 426 Auditorium Road, East Lansing 48824, Michigan, United States of America
[*]To whom correspondence should be addressed: ahz@du.se
FOR PUBLISHER ONLY Received on Date Month Year; accepted on Date Month Year

Abstract

Understanding the balance between evolvability and mutational robustness is crucial for exploring adaptation in complex fitness landscapes. This study examines how heterogeneous populations adapt under varying mutation rates (μ) and fitness landscape ruggedness (K), emphasizing their distinct starting conditions in valleys, slopes (low prominence), or peaks (high prominence). Using an extended NK model, we simulate populations capable of initiating anywhere in the landscape. Our findings reveal that starting positions strongly influence whether robustness or evolvability is advantageous. Populations beginning in low-prominence regions (valleys and slopes) exhibited high levels of epistasis (ϵ) and pleiotropy (π), enhancing evolvability and enabling exploration of the fitness landscape. In contrast, populations starting in high-prominence regions (peaks) reduced ϵ and π, prioritizing robustness to maintain stability against mutations. This study highlights the role of starting conditions in shaping evolutionary trajectories, offering insights into the interplay between evolvability and robustness.

Key words: epistasis; pleiotropy; fitness landscape; evolvability; robustness

Introduction

Mutational robustness and evolvability are fundamental aspects of evolution. The interplay between these two factors presents a complex and often paradoxical challenge in evolutionary biology [39, 32]. Mutational robustness refers to an organism's ability to maintain functionality despite genetic mutations, thereby preserving phenotypic stability [21, 39]. Conversely, evolvability denotes the potential to generate phenotypic variation, facilitating adaptation to changing environments [26]. The coexistence of these traits appears contradictory: robustness implies resistance to change, while evolvability necessitates variability [36, 11].

A key example of this paradox is observed in the "Survival of the Flattest" (SoF) phenomenon. A population at a high but steep fitness peak becomes out-competed by another population at a lower but flatter peak [41]. The reason for that lies in the population-level robustness [1] provided by the flatter peak.

While the population at the higher peak is more likely to produce mutants that "fall" from the peak due to the steep fitness landscape, the flatter peak accommodates a broader range of genotypes with minimal loss in fitness. This difference is sufficient for the survival of the population at the flatter peak.

Beyond simply occupying a steeper or flatter fitness peak, which determines the mutational neighborhood and thus the effect of mutations, organisms can modulate their inherent robustness against mutations to prevent changes to their phenotype [4, 34]. This means they are not merely passive participants shaped by the fitness landscape; rather, they can adjust their genetic architecture to buffer against mutations when phenotypic stability is advantageous or, conversely, to amplify the effects of mutations, allowing for greater exploration of the fitness landscape. Thus, a central question in evolutionary biology emerges: what mechanisms facilitate organisms to dynamically adjust between stabilizing their

---

[1] population-level robustness: refers to the ability of a population to maintain overall viability, fitness, or function

despite genetic mutations, environmental changes, or other perturbations

phenotype, and increasing phenotypic variation, thereby optimizing the need for both robustness and adaptability in response to changing environments?

We, and others, have identified epistasis (ϵ) and pleiotropy (π) to facilitate robustness and evolvability [7, 28, 11, 33, 17]. Epistasis refers to interactions between different genetic loci, where the effect of one gene mutation depends on the presence of mutations in other genes. This interaction can enhance robustness by buffering the effects of individual mutations, thereby stabilizing phenotypic traits despite genetic variability [31]. For instance, organisms with strong epistatic interactions can maintain functionality even with multiple mutations, effectively "hiding" the impact of deleterious changes. However, this stabilizing role can also limit evolvability by reducing the impact of mutations that might otherwise lead to beneficial adaptations [27]. Thus, epistasis can simultaneously support robustness while constraining the organism's ability to generate novel phenotypic variations [12].

Pleiotropy, on the other hand, occurs when a single gene influences multiple traits. This interconnectedness can contribute to mutational robustness by ensuring that changes in one gene lead to coordinated adjustments across multiple traits, thereby preserving overall functionality [35]. However, the downside is that mutations beneficial to one trait can have detrimental effects on others, thereby limiting the organism's ability to adapt optimally [7]. Lastly, a mutation in a pleiotropic gene can have multiple effects, as the pleiotropic gene is involved in many traits. As such, a mutation can now change the phenotype more drastically, leading to an increase in evolvability, as it allows the faster exploration of the fitness landscape [29, 35].

However, investigating these interactions in natural systems is challenging due to the intricacies of genetic networks and the limitations of experimental manipulation. To address this, computational models provide a powerful alternative, allowing researchers to systematically explore how varying mutation rates and fitness landscapes influence robustness and evolvability [40, 22]. In our recent extension of the NK fitness landscape model, we utilized a matrix-based encoding to map genetic loci to phenotypic traits, enabling a detailed investigation of how ϵ, π, and landscape ruggedness (K) affect adaptive potential [15].

Our previous work [17] explored how reducing ϵ, π can prevent populations from shifting toward flatter peaks under high mutation pressures. By minimizing genetic interactions, we showed that populations could maintain their fitness advantage despite elevated mutation rates. However, this research focused on populations already positioned on fitness peaks. Extending this, we have investigated how populations transition from lower to higher fitness peaks under different mutation rates, emphasizing the role of robustness in stabilizing populations once they reach peaks [18].

However, the focus so far has been on understanding how populations stabilize on a fitness peak or transition from a secondary peak to a primary peak. While these insights are valuable, they do not fully address the intricate balance between evolvability and robustness within the complex fitness landscapes. Populations can exist at any position within a fitness landscape–be it in valleys, intermediate plateaus, or random slopes–and their adaptive strategies from these starting points remain less explored. Specifically, we are interested in the conditions under which evolvability becomes advantageous over robustness, or vice versa.

Evolvability and robustness are highly context-dependent, shaped by an organism's position within the fitness landscape–whether on a peak, slope, or valley. On high-fitness peaks, robustness can limit adaptability by buffering against mutations, maintaining stability but reducing evolutionary potential. Because only a narrow range of genetic changes is tolerated without altering phenotype, adaptation to new environments is constrained [10, 34]. However, some argue that robustness can enhance long-term evolvability by allowing populations to accumulate neutral or slightly deleterious mutations, which may facilitate adaptive shifts when environmental conditions change [3, 13].

Organisms with moderate fitness (at slopes) are stable enough to resist harmful mutations but still flexible enough to accumulate beneficial ones. This allows them to maintain both genetic stability and the potential for adaptation [4]. However, some studies suggest that over time, selection on slopes may favor robustness, reducing adaptability unless external pressures actively maintain genetic exploration [34, 25].

In contrast, low-fitness valleys favor high evolvability, as populations tolerate a wide range of mutations, increasing the likelihood of discovering beneficial genetic variants despite initial fitness disadvantages [11, 2]. Yet, excessive mutational exploration can be detrimental, leading to genetic drift and a higher risk of extinction before beneficial mutations arise. In rugged landscapes, populations may struggle to escape valleys if adaptive paths require crossing multiple deleterious steps, limiting their evolutionary trajectory [37].

The relative advantages of robustness and evolvability depend on where a population is in its evolutionary trajectory. In the early stages, populations in low-fitness valleys benefit from high evolvability, as exploring a broad range of mutations increases their chances of adaptation. As fitness improves and populations move toward slopes or peaks, selection may begin to favor robustness to preserve advantageous traits. Whether this shift ultimately leads to robustness dominating or evolvability remaining crucial for long-term adaptability, however, remains unanswered.

Building on this foundation, our study aims to address the gap by investigating how the starting location of a population within a fitness landscape – whether in valleys, intermediate plateaus, or on high peaks–determines under what circumstances robustness or evolvability becomes advantageous. By simulating varying mutation rates (μ, ranging from 0.0001 to 0.05), landscape ruggedness (characterized by K), and initial population placements, our computational model provides new insights into how populations navigate fitness landscapes.

To explore these dynamics, we consider the role of population heterogeneity in navigating fitness landscapes. Natural systems often experience stochastic events such as bottlenecks or founder effects[20, 14], which dramatically alter population size and genetic composition. Bottleneck effects occur when population size is drastically reduced due to events like natural disasters, disease outbreaks, or intense selection pressures [20]. This reduction reshuffles genetic diversity and effectively resets the population's position within the fitness landscape. Similarly, founder effects arise when small subsets of populations establish themselves in novel environments, leading to unique genetic configurations and random starting locations in the fitness landscape [14, 30].

These scenarios highlight the importance of heterogeneity, as diverse genetic compositions provide the flexibility needed for populations to adapt to various positions in the fitness

landscape. We thus, incorporate heterogeneous populations into our computational model (see Material and Methods) to investigate how they leverage robustness and evolvability to traverse and adapt within rugged fitness landscapes. Recognizing the context-dependent nature of these traits, our study seeks to uncover the dynamics driving successful adaptation, highlighting the specific roles of robustness and evolvability across different evolutionary scenarios.

Materials and Methods

Extended NK Model

Our computational model is based on the standard NK model framework [8], which contains a circular sequence of N binary loci. In this model, each locus interacts with K adjacent loci to contribute to the organism's overall fitness. The model assigns a fitness value to each gene configuration via a lookup table that contains random values. For example, with $K = 1$, each gene is defined by the sequence of two adjacent loci, resulting in four possible binary configurations (00, 01, 10, 11).

As K increases, each gene is influenced by more neighboring loci, increasing the size of the fitness table. Since sites are overlapping, single mutations are affecting more genes, which leads to more changes in fitness values creating a more rugged fitness landscape with multiple peaks and valleys. Unlike the traditional approach of averaging fitness components, we calculate the organism's fitness (W) using the geometric mean of all N fitness values (see Equation 1), which has been shown to provide realistic evolutionary outcomes [38, 24, 23]:

$$W = \left(\prod_{i=1}^{N} W_i \right)^{1/N} \tag{1}$$

Genotype-to-Phenotype and Initial Starting Conditions

The main difference between the extended NK model [15] used here and the standard NK model [8] lies in the use of an indirect encoding mechanism. While fitness is still calculated based on the traits of the organism defined by its phenotype, the extended NK model introduces a genome that indirectly encodes the phenotype. This allows for more complex mappings between the genotype and phenotype, enabling the exploration of robustness and evolvability in greater detail.

In this experiment, we utilized the extended NK model to investigate the effects of mutational robustness and evolvability across 21 distinct starting conditions in the fitness landscape, corresponding to different configurations of neighboring fitness values. The study examines how mutation rates (μ), landscape ruggedness (K), and these starting conditions influence evolutionary outcomes.

A heterogeneous population of 100 organisms was initialized for each starting condition, with each organism represented by a genotype vector and a corresponding phenotype vector. These starting conditions were determined based on the number of neighboring genotypes with higher or lower fitness values. In each generation, mutations were introduced to both the genotype values (G) and the genome-to-phenotype mapping matrix (M)(see figure 2) at the specified mutation rate. These mutations allowed the population to explore the fitness landscape by introducing genetic variation. Selection was performed using a roulette-wheel mechanism[6], where organisms with higher fitness values had a greater probability of reproducing and passing on their genome to the next generation.

Prominence Analysis in the extended NK model

In the context of our study, prominence is a measure of how elevated a genotype is compared to its neighbors. It is calculated as the number of neighboring genotypes that have lower fitness than the focal genotype. A high prominence value indicates that the genotype is situated on a pronounced peak, surrounded by genotypes of lower fitness. Conversely, a low prominence value corresponds to genotypes that are closer to valleys, where most neighboring genotypes have higher fitness.

The figure (see Figure 1 Panel A.) illustrates how the distribution of prominence values changes with varying ruggedness levels (K). The parameter K controls the ruggedness of the fitness landscape: as K increases, the number of fitness peaks increases, creating more valleys and dips between the peaks. For lower K values, the landscape is smoother, and the highest peak corresponds to the highest fitness (1.0). But, as K increases, regions of high prominence are less common, and peaks tend to have lower fitness compared to the highest peak in smoother landscapes (see figure1 Panel B.). This pattern shows that as the landscape becomes more rugged, evolution becomes more constrained, with fewer accessible pathways for populations to climb fitness peaks.

Indirect Encoding

To isolate the effects of epistasis (ϵ) and pleiotropy (π) from landscape ruggedness (K), we extended the standard NK model by separating the genotype from the phenotype. In our model, each locus in the genotype vector takes a continuous value within the range $[-1, 1]$ rather than binary values. This extension allows us to apply a mapping matrix that transforms these genotype values into a phenotype vector, which is subsequently converted to binary for fitness calculations.

We use a mapping matrix, M, of dimension $N \times N$, with values in the range of $[-1, 1]$ [15, 16, 17]. This matrix defines the interactions between genes, influencing each phenotypic trait according to the genotype values. The phenotype vector, P, is calculated as the dot product of the genotype vector G and the mapping matrix M (see Equation 2):

$$P = G \cdot M \tag{2}$$

Each element of the phenotype vector, representing a trait, is subsequently converted to binary–values greater than 0 are assigned a binary value of 1.0, while values below or equal to 0 are assigned 0.0. This phenotype vector (P), is then treated as the trait vector in the standard NK model, allowing fitness to be calculated based on the value of K (see Figure2). However, unlike in the standard NK model, the degree to which genes can interact is now freely evolvable by changing the expression values G and the weights of the interaction matrix M, independent of K.

Quantifying Epistasis and Pleiotropy

In our extended model[15], we have separated genetic interactions to allow traits to evolve independently. This separation requires distinct measurements for epistasis (ϵ) and pleiotropy (π).

To achieve this, we calculate an interaction matrix (IM), which identifies how each gene influences specific traits. For an organism with N genes and N traits, the IM is an $N \times N$ matrix[15]. We determine whether each gene i affects a trait j by altering the expression value G_i between -1.0 and 1.0. Using these modified values, we construct the phenotype $P = G \times M$
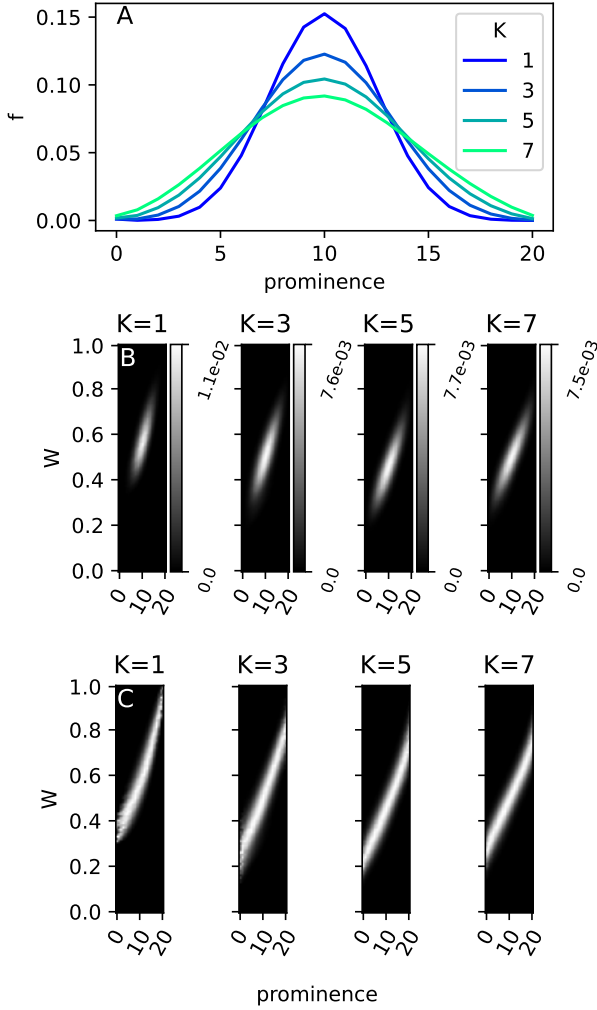
Fig. 1. The figure illustrates the impact of landscape ruggedness ($K = 1, 3, 5, 7$) on fitness prominence distributions. (A) The top plot shows the mean prominence values of fitness peaks for different ruggedness levels. Prominence, normalized by the total size of the landscape, decreases as ruggedness (K) increases. The colors represent different K values, with higher K leading to smoother, lower peaks. (B) The middle panel displays the probability density function of the prominence as a grayscale heatmap across all levels of ruggedness (x-axis). The adjacent color legend shows the maximum value at the top, scaling from white for that maximum to black for a density of 0.0. (C) The bottom panel displays grayscale heat maps of prominence distributions for each ruggedness level, but normalized along each column because low and high prominences are less often found than intermediary prominences. These heatmaps highlight how increasing K disperses fitness prominence, leading to flatter and more widespread peak distributions.

and check if changing gene i impacts trait j. If it does, $IM_{i,j}$ is marked with a value of 1; if not, it is assigned 0. The row sums in this matrix indicate the number of traits influenced by each gene (π), while the column sums reflect the number of genes affecting each trait (ϵ). This gene-to-trait mapping is governed by the mapping matrix M, which evolves over time. Due to random mutations, directly comparing ϵ and π vectors across organisms can be challenging. We resolve this by sorting these vectors in ascending order[15], allowing us to analyze their overall distribution at an organismal level rather than focusing on specific values for individual genes.
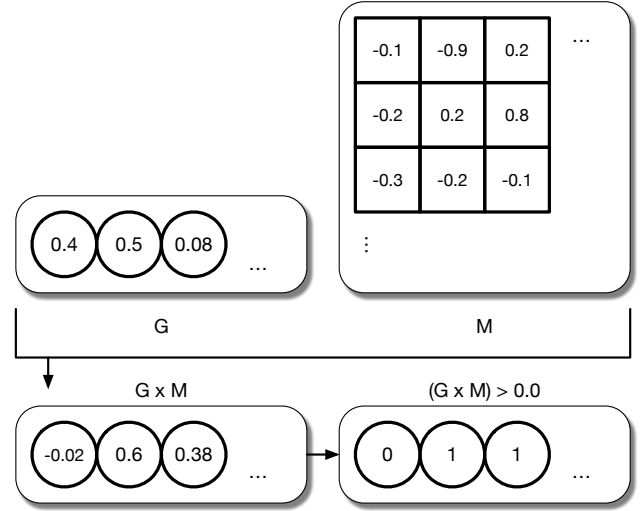

Fig. 2. In the indirect encoding method, each organism has a genome composed of N genes. Each gene, represented as a circle, is characterized by a specific value and an associated vector of length N. These gene vectors collectively create a matrix (M), while the gene values form a vector G. The phenotype vector is then produced by taking the dot product of G with M, after which it is discretized so that each component corresponds to one of the N binary traits of the phenotype. Mutations can alter any values within the genome, allowing the interaction between G and M to indirectly shape the phenotype. For direct encoding, M would instead be a fixed identity matrix.

We then compare the observed ϵ and π distributions to a baseline random expectation, represented by the black line (see Figure 3). This expectation is calculated from 1 000 mapping matrices with randomly generated values between -1 and 1. Values below this baseline indicate lower-than-expected ϵ and π, while values above suggest higher-than-expected levels [15].

To quantify these differences from the expected ϵ and π values, we calculate the area above and below the baseline curves, denoted a and b respectively. The discrepancy Δ, defined as the difference between these two areas, represents how much the observed values deviate from the random expectation.

Generating Heterogeneous Populations with Varied Epistasis and Pleiotropy

To generate a heterogeneous population with varied levels of ϵ (epistasis) and π (pleiotropy), a **hill-climber optimization** approach is employed. This method compares an organism with its mutant counterpart, maintaining an identical phenotype but with varied levels of ϵ or π. We achieve this by selecting organisms with either increased or decreased levels of ϵ or π. Depending on our goal, we can maximize or minimize ϵ and π, or choose one randomly to emulate genetic drift.

Each iteration introduces a random mutation until a mutant with the same phenotype as the original is generated. If this mutant shows a difference in ϵ or π consistent with the desired direction (higher or lower), it is retained. This process is repeated for up to 1 000 iterations, ensuring the phenotype remains unchanged, thereby fixing the organism's position within the fitness landscape.

Once this process is completed, a heterogeneous population of 100 organisms is created by selecting half of the organisms from the pool of organisms with increased values of ϵ and π,
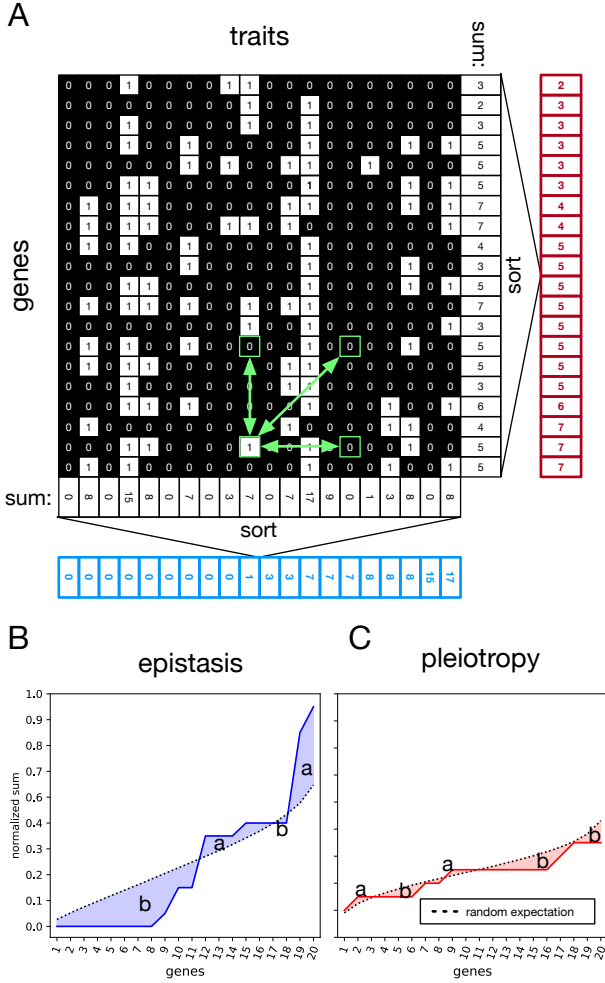
A

B epistasis **C** pleiotropy

Fig. 3. To figure illustrates how epistasis (ϵ) and pleiotropy (π) are calculated as deviations from expected interactions, consider the following. Panel A depicts an organism's interaction matrix (IM), where the rows represent $N = 20$ genes and the columns represent $N = 20$ traits. A value of 0 signifies no interaction between a gene and a trait, while a value of 1 indicates that the gene has an influence on the trait. Pleiotropy is determined by summing each row (number of traits influenced by a gene), and epistasis is calculated by summing each column (number of genes influencing a trait). These sums are organized into vectors and sorted in ascending order. In Panel B, the normalized epistasis vector is represented by a solid blue line, while the random expectation is shown as a dashed black line. The difference between these two is illustrated as a blue area, with portions above the expectation labeled a and those below labeled b. The same approach is shown in Panel C for pleiotropy, where the results are displayed in red. The total difference between the measured values and the expectation is calculated as $a - b$. Typically, the observed values (solid blue or red lines) remain consistently on one side of the expected line (dashed black line), but this method also accommodates scenarios where the observed values fluctuate around the expectation. Changes to the interaction matrix, such as the purple arrows in Panel A indicating swapped interactions, can have different impacts on pleiotropy and epistasis. For example, swapping interactions horizontally alters epistasis, vertical swaps affect pleiotropy, and diagonal changes impact both. Similarly, adding or removing interactions at specific locations modifies both pleiotropy and epistasis simultaneously. This method ensures that the differences between observed and expected interactions are measured accurately.

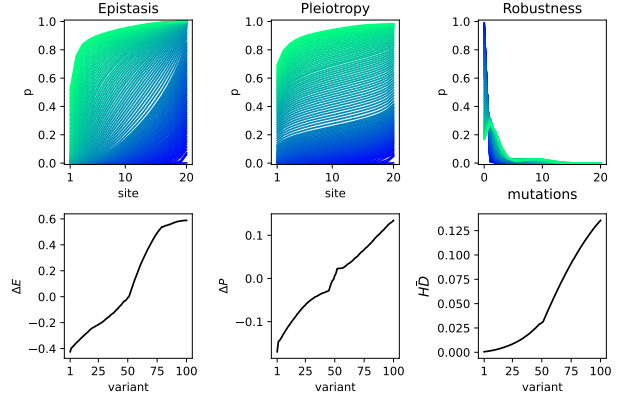and the other half from the pool with decreased values. If the

Fig. 4. The figure illustrates the probability distributions of epistasis(left), pleiotropy (middle), and mutational robustness (right) in a fitness landscape. The top row shows individual simulation trajectories, where the x-axis represents prominence (epistasis, pleiotropy) or mutations (robustness) and the y-axis indicates probability p. The bottom row presents change in $\Delta\epsilon$ $\Delta\pi$ and $\overline{HD}$ across 100 organisms (black line).

number of organisms is insufficient, duplicates are added, and if there are too many, organisms are removed.

This heterogeneous composition enables a comprehensive analysis of how different genetic interaction levels impact mutational robustness and evolutionary dynamics across a diverse population.

Mapping the Fitness Landscape

To explore how populations adapt and evolve in diverse fitness landscapes, we must carefully sample possible starting conditions populations that just arrived in a new evolutionary context find themselves in. While organisms could theoretically start from any position within the landscape, sampling all potential positions is computationally infeasible. Instead, we focus on using **prominences**, which describes the degree to which a location in the landscape is a peak. We measure prominence as the number of neighboring genotypes with lower fitness values. With $N = 20$, there are $N + 1$ possible prominence levels, ranging from 0 (deepest dips) to 20 (highest peaks). However, each location, given its prominence, can have different fitness values. While there is, for example, only one highest peak with a prominence of $N = 20$, other peaks have lower fitness values. To account for the diversity of fitness values for different prominences, we select 10 start locations for each prominence level, equally spaced across the possible fitness ranges at each prominence level; ensuring a balanced and diverse representation of the fitness landscape.

Specifically, we compute the prominence of every possible 2^N phenotypic start locations, and its fitness values, and then select the $(N + 1) \times 10$ start locations given that information. For each of these start locations, straddling the entire fitness landscape, a population of 100 organisms is created by enriching the diversity for ϵ and π as described above.

Evolution of Populations

After initializing the population with 100 organisms, all sharing the same phenotype but with varied levels of ϵ and π, the simulation employs a genetic algorithm to evolve the population over multiple generations. In each generation, the fitness of each organism is assessed based on its phenotype's alignment

with a predefined fitness landscape. A roulette-wheel selection method is then used to favor organisms with higher fitness for reproduction.

Mutations occur, introducing changes to both the genotype (G) and the mapping matrix (M), which can alter ϵ and π. These mutations enable the population to explore the fitness landscape, driving adaptation or genetic drift. The process concludes when the population reaches a genetic sweep. When running controls without fitness differences to model adaptation under drift the experiment was stopped after 200 generations – which is the expected fixation time for a random trait under drift[9]

Quantifying Mutational Robustness ($\overline{HD}$)

In this study, we assess mutational robustness by examining the average impact that mutations have on an organism's phenotype. Specifically, mutational robustness is considered lower when mutations cause a larger effect on the phenotype, facilitating the exploration of new regions in the fitness landscape.

To quantify this impact, we calculate the average Hamming distance ($\overline{HD}$) between the original organism and its mutants. Each organism generates $10\,000$ offspring, with a point mutation rate of 0.01 applied to both the expression levels and interaction weights (M) in the genotype. For each of these $10\,000$ mutants, the phenotype is determined, and the Hamming distances are measured and averaged to obtain $\overline{HD}$.

Although HD is measured abstractly within our genotype to phenotype model, it is conceptually analogous to biological measures of robustness used in molecular evolution, such as protein or RNA stability [1, 32], where robustness is assessed by the degree to which mutations preserve structural or functional traits.

When correlating the mean epistatic and pleiotropic effects to $\overline{HD}$, we observed that randomly generated organisms exhibit only minor variations in ϵ and π. To address this, we employed a hill-climbing optimization strategy that attempts up to $1\,000$ iterations to generate mutants that maintain the same phenotype but have either increased or decreased levels of ϵ and π[17] (see Figure 4)

Simulation of Genetic Drift

The process begins by creating a root organism with a specific phenotype, which serves as the ancestor for the population. From this organism, a population of 100 organisms is created, ensuring that all organisms initially share the same phenotype but have varied levels of ϵ) and π. The simulation then evolves the population purely through genetic drift until a genetic sweep occurs. During this phase, organisms reproduce randomly to form new generations without fitness-based selection, allowing genetic drift to drive changes in the population.

On average, a genetic sweep—where the entire population descends from a single lineage—occurs within approximately 200 generations, equivalent to 2× the population size (100). To standardize results, the simulation length is normalized to 200 generations, ensuring consistency across experiments [9].

Details of our computational model

The code for the computational model can be found here . Populations were initialized with 100 organisms, each having distinct genotypes but sharing the same initial phenotype, derived from a root organism (heterogeneous population). At every generation, the entire population was replaced with a new one based on fitness values using roulette wheel selection. This stochastic selection method ensured that individuals with higher fitness were more likely to reproduce, while less fit individuals retained a small chance to pass their traits to the next generation.

The evolutionary process was strictly asexual, with no recombination or horizontal gene transfer. Mutations occurred probabilistically based on the mutation rate (μ), affecting both the genome and the interaction matrix, which indirectly influenced the phenotype. The population size of 100 was maintained constant throughout all simulations. Fitness landscapes with varying ruggedness ($K = 1, 3, 5, 7$) were used, enabling exploration of how populations adapted under different levels of landscape complexity.

Results

Previous research has provided two key insights that shaped our hypothesis. First, populations already positioned on fitness peak benefit from mutational robustness, particularly under high mutation rates [17]. Robustness minimizes the impact of deleterious mutations, allowing populations to maintain phenotypic stability in the face of genetic changes. Second, while evolvability facilitates crossing fitness valleys and reaching higher fitness peaks, long-term stability at these peaks requires robustness. Highly evolvable pioneers often find themselves out-competed by more robust organisms once the population stabilizes on a peak, as robustness ensures greater persistence under continued mutational pressure [18].

While these two principles make general suggestions about how organisms behave when sitting at a peak or how they cross a valley, we do not know how a population that could have been placed anywhere in the fitness landscape adapts, and if mutational robustness or evolvability conveys a greater advantage. Also, the shape of the fitness landscape (ruggedness) and with it the distribution of fitness values and prominences must influence said evolutionary success and type of adaptation.

To address this, we define evolutionary success as happening at the moment when any organism from the initial population sweeps the population and becomes the ancestor of all future generations. For example, if a population starts at the lowest valley, does the most evolvable organism dominate and sweep the population, or does an organism with high mutational robustness become the ancestral lineage? Starting at an intermediate or high peak might yield different outcomes. To explore these dynamics, we tested the roles of robustness and evolvability across all possible starting conditions in the fitness landscape, ranging from the lowest valleys to the highest peaks and intermediate slopes.

While our attention is on mutational robustness and evolvability, we also showed that they are controlled by epistasis (ϵ) and pleiotropy (π). Therefore, we will monitor all four factors in the following experiments.

We conducted simulations using populations initialized at diverse starting positions (see Materials and Methods) across the fitness landscape, ranging from valleys (low fitness) to peaks (high fitness). By varying mutation rates (μ) 0.0001. 0.001, and 0.01 and ruggedness levels ($K = 1, 3, 5, 7$), we explored how these factors influence the relative advantages of evolvability and robustness. We ignore $K = 0$, as that is an entirely flat landscape with a single peak that can be trivially
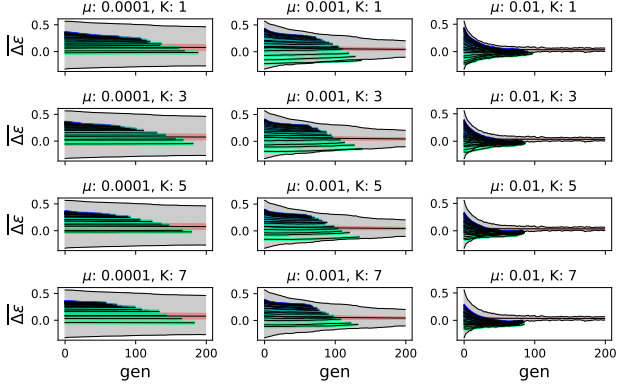
Fig. 5. Temporal dynamics of epistasis (ϵ) across evolutionary simulations for populations initialized at different prominence levels in the fitness landscape. Each subplot represents a specific combination of landscape ruggedness ($K = 1, 3, 5, 7$) and mutation rate ($\mu = 0.0001, 0.001, 0.01$). The x-axis indicates generation number (up to 200), while the y-axis shows changes in epistasis ($\Delta\epsilon$) relative to baseline. Black lines represent the mean $\Delta\epsilon$ trajectories from different starting positions. The blue-green shadow regions indicate the 95% confidence interval. The long black line depicts the average trajectory of $\Delta\epsilon$ across simulation replicates, with its 95% confidence interval represented as a red shadow, and the shaded region encompasses the minimum and maximum observed values, illustrating the range of evolutionary outcomes. This figure captures how different levels of landscape ruggedness and mutation rate shape the degree of epistasis over time.
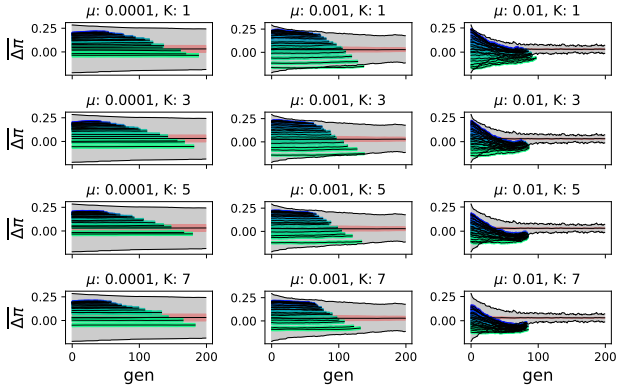

Fig. 6. Temporal dynamics of pleiotropy (π) during evolution from diverse starting positions in the fitness landscape. Each subplot corresponds to a specific combination of landscape ruggedness ($K = 1, 3, 5, 7$) and mutation rate ($\mu = 0.0001, 0.001, 0.01$). The x-axis indicates generation number, while the y-axis shows changes in pleiotropy ($\Delta\pi$) relative to initial conditions. Black lines represent the mean $\Delta\pi$ trajectories from different starting positions. The blue-green shadow regions indicate the 95% confidence interval. Long black line denotes the mean $\Delta\pi$ across replicates, with its 95% confidence interval represented as red shadow. Gray shaded regions outline the full range (minimum to maximum) of drift outcomes across simulations. This figure highlights how the interplay between mutation rate and landscape ruggedness influences the evolution of pleiotropy over time.

reached. Tracking changes in epistasis (ϵ) and pleiotropy (π) over generations provided insight into the genetic interactions driving adaptability and stability. Control simulations without selection pressure were also performed to isolate the effects of random drift and ensure the observed dynamics were driven by selection.

Populations starting in low-to-medium prominence regions (prominence levels 0 to 17) exhibit high levels of epistasis (ϵ) and pleiotropy (π), indicative of high evolvability (see Figures 5, 6). In contrast, populations starting in high-prominence regions (prominence levels 18 to 20), such as fitness peaks, display lower levels of ϵ and π. This reduction in evolvability corresponds to greater mutational robustness.

However, over the course of evolution, populations experiencing high mutation rates tend to become more robust. If their mutational robustness is already high at the outset (low ϵ and π), it may decrease slightly over time but remain below the random expectation. At low mutation rates, populations remain unaffected.

We measured mutational robustness using the average Hamming distance ($\overline{HD}$), which quantifies the phenotypic impact of mutations. Higher $\overline{HD}$ indicates lower mutational robustness, while lower $\overline{HD}$ reflects greater phenotypic stability. Consistent with prior findings [18], $\overline{HD}$ correlates with epistasis (ϵ) and pleiotropy (π): higher ϵ and π increase $\overline{HD}$, facilitating exploration but reducing robustness.

Populations starting in low-to-medium prominence regions exhibit higher $\overline{HD}$, highlighting their reliance on evolvability for exploration. In contrast, populations beginning on peaks maintained lower $\overline{HD}$, especially under high mutation rates, favoring robustness (see Figure 7).

The ruggedness of the landscape (K) does not appear to influence the advantage of highly evolvable organisms starting in valleys or highly robust organisms starting on peaks. Similarly, adaptations that adjust $\overline{HD}$ in high mutation rate environments remain unaffected by landscape ruggedness (K). However, this does not mean that K does not impact the results. In landscapes with higher K, which statistically feature more peaks and valleys, the differences in robustness and evolvability become more significant during the sweeping phase. Nonetheless, these differences do not extend to later stages of adaptation (see Figure 1 to see an illustration about landscapes with higher K having more valleys and peaks).

Populations starting in low-to-medium prominence benefit from evolvability, with low mutational robustness enabling rapid adaptation and exploration. Conversely, populations starting from locations with higher prominence rely on robustness to maintain stability.

Statistical Verification

The above observations can also be supported with statistical evidence. To investigate whether the observed evolutionary trajectories of mutational robustness ($\overline{HD}$) significantly deviate from what would be expected under neutral drift, we performed a Kolmogorov-Smirnov (KS) test [5]. This test compared the distributions of HD at the start and end points of the observed evolutionary trajectories with conditions under which organisms only experienced drift. By doing so, we assessed whether the deviations in $\overline{HD}$ were statistically significant and indicative of adaptive responses rather than random drift.

Figure 8 illustrates the results of this analysis. The black solid line represents the distribution of the starting points, while the red solid line represents the distribution of the endpoints. The dotted red line corresponds to the drift baseline.

The p-values are consistently low, ranging from 1.0264×10^{-19} to 2.5210×10^{-8}, across all tested mutation rates ($\mu = 0.0001, 0.001, 0.01$) and landscape ruggedness levels ($K = 1, 3, 5, 7$). This confirms that the changes in $\overline{HD}$ observed
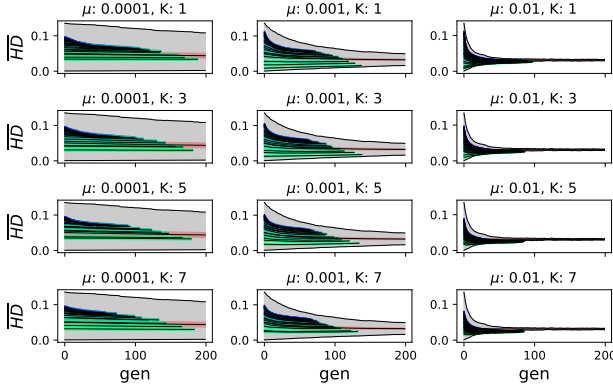
Fig. 7. Evolutionary trajectories of mutational robustness, measured as average Hamming distance ($\overline{HD}$), across generations for populations initialized at varying positions within the fitness landscape. Each subplot presents a specific combination of landscape ruggedness ($K = 1, 3, 5, 7$) and mutation rate ($\mu = 0.0001, 0.001, 0.01$). The x-axis denotes the number of generations, and the y-axis displays $\overline{HD}$, where higher values indicate lower robustness (greater phenotypic change upon mutation). Black lines represent the mean $\overline{HD}$ trajectories from different starting positions. The blue-green shadow regions indicate the 95% confidence interval. The Long black line represents the mean $\overline{HD}$ across replicate simulations, with its 95% confidence interval represented as a red shadow, and shaded regions show the full observed range. This figure demonstrates how mutation rate and landscape ruggedness jointly shape the evolution of mutational robustness.

Fig. 8. Comparison of mutational robustness distributions at the start and end of evolution, alongside a neutral drift baseline, across different evolutionary conditions. Each subplot represents a unique combination of landscape ruggedness ($K = 1, 3, 5, 7$; rows) and mutation rate ($\mu = 0.0001, 0.001, 0.01$; columns). The black solid line shows the average Hamming distances ($\overline{HD}$) at initialization, the red solid line shows the same but after evolution, and the black dotted line represents the drift-only control. Across all conditions, the evolved distributions significantly diverge from the drift baseline, confirming that changes in robustness arise from adaptive responses rather than stochastic effects alone.

during evolution are driven by adaptation to selective pressures and not merely the result of neutral drift.

Adaptation of Mutational Robustness over the Course of Evolution

To understand how mutational robustness (HD) adapts during evolution, we analyzed its relationship with prominence under varying mutation rates (μ) and landscape ruggedness (K). As organisms evolve, they typically increase their fitness and occupy higher peaks. To test whether mutational robustness increases as a result of adaptation, we plotted the robustness of organisms along their line of descent (LOD) against their prominence. Our results confirm an overall inverse correlation between $\overline{HD}$ and prominence (dashed line) across all data (see Figure 9).

When differentiating by starting conditions, we observed that under high and intermediate mutation rates, mutational robustness consistently increases with prominence. At low mutation rates, this trend holds for most start conditions but is weaker for populations starting in low-prominence regions, where evolvability appears more critical than robustness. Lastly, the ruggedness of the landscape (K) does not seem to directly influence this pattern beyond the aforementioned bias in start conditions.

Therefore, as long as the mutation rate is high enough, organisms generally tend to evolve towards having more mutational robustness as they gain more fitness. In other words, at the start of evolution, having a high evolvability is more advantageous, suggesting a higher need for valley crossing, which becomes less important as organisms settle at peaks where high mutational robustness provides the highest advantage.
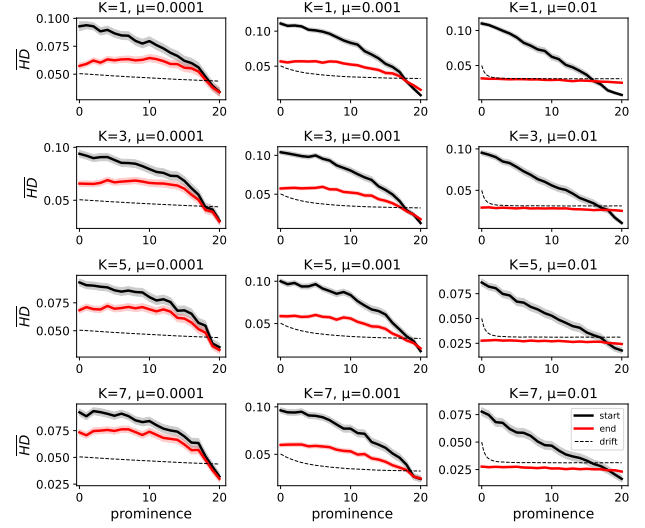
Discussion

In this study, we explored the complex interplay between evolvability and mutational robustness within fitness landscapes of varying ruggedness. By utilizing an extended NK model with indirect encoding, we examined how populations navigate fitness landscapes starting from valleys, slopes, or peaks under varying mutation rates and levels of landscape ruggedness. Our findings confirm that the evolutionary trajectory of a population is critically shaped by its starting position in the landscape and the dynamic trade-offs between evolvability and robustness.

Generally speaking, the fitness landscape correlates low fitnesses with valleys and thus low prominence, while areas in the fitness landscape that have higher fitness naturally have higher prominence.

The results demonstrate that populations starting in valleys or on slopes favor evolvability as a mechanism for exploration. Elevated levels of epistasis (ϵ) and pleiotropy (π) in these populations reflect an increased genetic interaction and a broad phenotypic impact of mutations, respectively. This adaptability allows populations to traverse rugged fitness landscapes, overcome valleys, and inhabit higher fitness peaks. However, as these populations ascend to fitness peaks, they undergo a transition: ϵ and π decrease, signifying a shift toward mutational robustness. This adaptation is particularly evident under high mutation rates, where stabilizing advantageous traits becomes essential for long-term survival.

Additionally, populations starting on fitness peaks exhibit higher levels of mutational robustness, reflected by lower Hamming distances (HD). This robustness allows populations to buffer against mutations, maintaining phenotypic stability and preserving their advantageous position on the peaks.
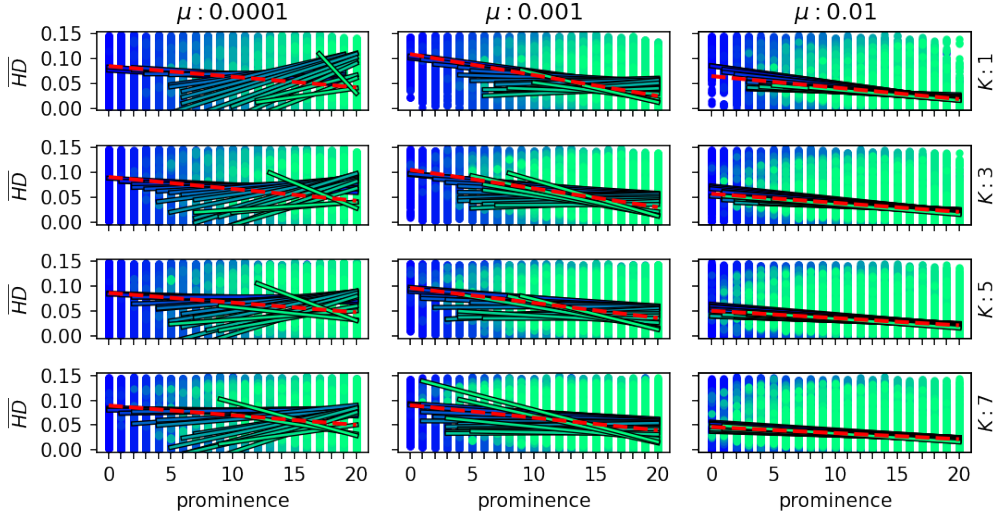
Fig. 9. Relationship between mutational robustness, measured by average Hamming distance ($\overline{HD}$), and prominence across evolutionary trajectories. The x-axis represents genotype prominence, and the y-axis indicates $\overline{HD}$, where lower values correspond to greater robustness. Each data point represents an organism along its line of descent (LOD), with colors denoting different starting conditions. The dashed line shows the overall inverse correlation: as prominence increases, $\overline{HD}$ tends to decrease, reflecting increased robustness. This trend is consistently observed under high and intermediate mutation rates (μ), while it is weaker under low mutation rates, especially for populations beginning in low-prominence regions. The landscape ruggedness parameter (K) does not strongly influence this pattern beyond its role in determining initial starting conditions.

High prominence, indicative of distinct and isolated fitness peaks, correlates with reduced HD, underscoring the adaptive advantage of robustness for populations at evolutionary endpoints. Notably, this robustness often comes at the cost of reduced evolvability.

The role of mutation rate (mu) further underscores the context-dependent nature of these dynamics. At low μ, populations rely heavily on evolvability to explore the fitness landscape, with reduced selective pressure on robustness. As μ increases, the cost of maintaining high evolvability rises, favoring the evolution of robustness. This observation aligns with previous studies showing that mutational robustness can be an essential factor in high-mutation environments[17].

While our study provides critical insights into the evolutionary dynamics of robustness and evolvability, it is limited by the assumptions of the computational model. The simplified representation of fitness landscapes and the focus on asexual reproduction do not fully capture the complexity of natural systems. Incorporating additional features such as recombination, gene regulatory networks, and environmental fluctuations could provide a more comprehensive understanding of these dynamics.

One limitation of our current model lies in its temporal scope. While we simulate population dynamics across hundreds of generations – sufficient to observe the transition from early exploration (evolvability-driven) to stabilization on fitness peaks (robustness-driven) – these timeframes are relatively short in evolutionary terms. As such, we may not capture long-term evolutionary dynamics, including repeated landscape shifts, secondary adaptations, or the cumulative effects of historical contingency. Given our model's framing around ecological scenarios like founder events (e.g., colonization of a new environment), our results are most applicable to initial phases of adaptation. Future work could expand the genotype space (larger N), introduce environmental change over time, or compare different evolutionary timescales to assess whether the early trade-offs we observe persist or evolve under different selective regimes.

Another important limitation of our current model is the assumption of a static fitness landscape. In natural systems, environmental fluctuations often drive adaptation and can reshape selective pressures over time. In static landscapes, prominence is a fixed structural property; however, under temporal change, the fitness of genotypes—and thus their prominence—can vary dynamically. This introduces new challenges for interpreting robustness and evolvability, particularly in terms of time-dependent selective regimes.

In ongoing work, we have extended the NK model to include explicit temporal changes in the fitness landscape, allowing us to examine how robustness and evolvability behave under varying rates of environmental fluctuation [19]. This new framework enables experiments on how quickly populations must adapt, and whether certain levels of robustness are favored under specific change dynamics. While these ideas are beyond the scope of the current study, they represent a natural and promising extension.

Our results suggest a relatively direct trade-off between mutational robustness and evolvability, mediated by epistasis () and pleiotropy (). However, this relationship is likely oversimplified. In biological systems, robustness can actually promote evolvability by enabling the accumulation of cryptic genetic variation that becomes advantageous in changing environments [13, 4, 2]. Furthermore, robustness can buffer phenotypes against deleterious mutations while allowing neutral or nearly-neutral variation to persist, enhancing future adaptive potential [32]. The relationship between robustness and evolvability is therefore context-dependent, influenced by environmental dynamics, mutation rates, and the structure of genotype–phenotype maps. For example, fluctuating environments have been shown to favor phenotypic plasticity and diversity, challenging the notion of a fixed trade-off [10]. In addition, modularity in gene-to-trait mappings

can decouple robustness and evolvability, allowing systems to be robust in some traits while remaining evolvable in others [35]. Future extensions of our model could integrate these features—temporal environmental change, modularity, and the potential for recombination—to explore how robustness and evolvability co-evolve under more biologically realistic conditions.

Our model does not incorporate recombination, horizontal gene transfer, or sexual reproduction, all of which play important roles in many biological systems. These mechanisms can substantially alter evolutionary dynamics by reshuffling genetic material, influencing standing variation, and shaping fitness landscapes. While our goal was to isolate the contributions of epistasis, pleiotropy, and mutational robustness in a simplified asexual framework, future work could expand the model to explore the role of recombinational robustness, dominance effects, and sexual selection. Such extensions would be especially relevant for studying diploid systems or organisms with complex life cycles, where multi-copy gene interactions and mating strategies add additional evolutionary pressures.

Conclusions

This research unveils the intricate relationship between evolvability and robustness in evolutionary systems. By quantifying genetic interactions and their phenotypic impacts through ϵ, π, and HD we demonstrate that these traits are not fixed but evolve in response to selective pressures and environmental challenges. Populations starting in valleys or slopes leverage evolvability to adapt and climb fitness peaks, while populations on peaks prioritize robustness to maintain their advantageous positions.

Our findings have implications beyond theoretical models, offering insights into real-world evolutionary processes. In microbial populations, for instance, the balance between evolvability and robustness can influence responses to antibiotics, where early adaptation requires flexibility, but long-term stability demands resilience. Similarly, evolutionary algorithms in computational contexts might benefit from adjusting levels of genetic interaction to optimize search strategies and stabilize solutions.

Or, for a more concrete example, imagine the transmission of a disease through a droplet filled with infectious particles varying in their mutational robustness and evolvability. Depending on their new environment, when starting at a valley, evolvability matters most, but when finding themselves at a peak, mutational robustness dominates. As such, one can not simply think of such an environmental change as a bottleneck, but whether the organisms find themselves in a valley, at a peak, or somewhere else in the new environment matters most.

This study is essential for understanding how populations adapt by balancing evolvability and robustness. In biological systems, this trade-off plays a key role in antibiotic resistance. Initially, bacteria rely on high evolvability to generate beneficial mutations that enable survival under antibiotic pressure. Once resistance is established, robustness becomes crucial to maintain the adaptation and prevent its loss due to harmful mutations.

A similar dynamic occurs in "island models", used in both biological evolution and computational optimization. In these models, isolated populations evolve independently before periodically exchanging genetic material. This study suggests that different "islands" may adopt distinct strategies–some favoring exploration through high evolvability, while others prioritize stability through robustness. Understanding this balance can inform both evolutionary biology and algorithm design, optimizing adaptability while ensuring long-term stability.

Future research should extend these findings by incorporating additional biological complexities and testing these principles in experimental evolution studies. By bridging theoretical and empirical approaches, we can deepen our understanding of how robustness and evolvability shape the adaptive potential of populations in diverse environments.

Data availability

The data underlying this article are available in [repository name, eg, the GenBank Nucleotide Database] at [URL], and can be accessed with [unique identifier, eg, accession number, deposition number].

References

1. Jesse D Bloom, Sy T Labthavikul, Christopher R Otey, and Frances H Arnold. Protein stability promotes evolvability. *Proceedings of the National Academy of Sciences*, 103(15):5869–5874, 2006.
2. Anton Crombach and Paulien Hogeweg. Evolution of evolvability in gene regulatory networks. *PLoS computational biology*, 4(7):e1000112, 2008.
3. Jeremy Draghi and Gunter P Wagner. The evolutionary dynamics of evolvability in a gene network model. *Journal of evolutionary biology*, 22(3):599–611, 2009.
4. Jeremy A Draghi, Todd L Parsons, Günter P Wagner, and Joshua B Plotkin. Mutational robustness can facilitate adaptation. *Nature*, 463(7279):353–355, 2010.
5. Guillaume J Filion. The signed kolmogorov-smirnov test: why it should not be used. *Gigascience*, 4(1):s13742–015, 2015.
6. David E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Publishing Company, Reading, MA, 1989.
7. Thomas F Hansen. Is modularity necessary for evolvability?: Remarks on the relationship between pleiotropy and evolvability. *Biosystems*, 69(2-3):83–94, 2003.
8. Stuart Kauffman and Simon Levin. Towards a general theory of adaptive walks on rugged landscapes. *Journal of theoretical Biology*, 128(1):11–45, 1987.

9. Motoo Kimura. *Population genetics, molecular evolution, and the neutral theory: selected papers.* University of Chicago Press, 1994.

10. Edo Kussell and Stanislas Leibler. Phenotypic diversity, population growth, and information in fluctuating environments. *Science*, 309(5743):2075–2078, 2005.

11. Richard E Lenski, Jeffrey E Barrick, and Charles Ofria. Balancing robustness and evolvability. *PLoS biology*, 4(12):e428, 2006.

12. Richard E Lenski, Charles Ofria, Travis C Collier, and Christoph Adami. Genome complexity, robustness and genetic interactions in digital organisms. *Nature*, 400(6745):661–664, 1999.

13. Joanna Masel and Meredith V Trotter. Robustness and evolvability. *Trends in Genetics*, 26(9):406–414, 2010.

14. Ernst Mayr et al. Change of genetic environment and evolution. 1954.

15. Priyanka Mehra and Arend Hintze. An extension to the nk fitness landscape model to study pleiotropy, epistasis, and ruggedness independently. In *2022 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1259–1267. IEEE, 2022.

16. Priyanka Mehra and Arend Hintze. Evolution of pleiotropy and epistasis in a gene regulatory network. In *Artificial Life Conference Proceedings 35*, volume 2023, page 95. MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info . . . , 2023.

17. Priyanka Mehra and Arend Hintze. Reducing epistasis and pleiotropy can avoid the survival of the flattest tragedy. *Biology*, 13(3):193, 2024.

18. Priyanka Mehra and Arend Hintze. The role of pleiotropy and epistasis on evolvability and robustness in a two-peak fitness landscape. *Biology*, 13(12):1003, 2024.

19. Priyanka Mehra and Arend Hintze. Continuous evolution in the nk treadmill model. *Artificial Life*, pages 1–20, 2025.

20. Masatoshi Nei, Takeo Maruyama, and Ranajit Chakraborty. The bottleneck effect and genetic variability in populations. *Evolution*, pages 1–10, 1975.

21. H Allen Orr. The rate of adaptation in asexuals. *Genetics*, 155(2):961–968, 2000.

22. H Allen Orr. Fitness and its role in evolutionary genetics. *Nature Reviews Genetics*, 10(8):531–539, 2009.

23. Bjørn Østman. Effects of epistasis and pleiotropy on fitness landscapes. *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*, pages 37–46, 2013.

24. Bjørn Østman, Arend Hintze, and Christoph Adami. Impact of epistasis and pleiotropy on evolutionary adaptation. *Proceedings of the Royal Society B: Biological Sciences*, 279(1727):247–256, 2012.

25. Joshua L Payne and Andreas Wagner. The robustness and evolvability of transcription factor binding sites. *Science*, 343(6173):875–877, 2014.

26. Massimo Pigliucci. Is evolvability evolvable? *Nature Reviews Genetics*, 9(1):75–82, 2008.

27. Frank J Poelwijk, Daniel J Kiviet, Daniel M Weinreich, and Sander J Tans. Empirical fitness landscapes reveal accessible evolutionary paths. *Nature*, 445(7126):383–386, 2007.

28. Rafael Sanjuán and Santiago F Elena. Epistasis correlates to genomic complexity. *Proceedings of the National Academy of Sciences*, 103(39):14402–14405, 2006.

29. Tanja Schwander and Olof Leimar. Genes as leaders and followers in evolution. *Trends in Ecology & Evolution*, 26(3):143–151, 2011.

30. Alan R Templeton. The theory of speciation via the founder principle. *Genetics*, 94(4):1011–1038, 1980.

31. Andreas Wagner. Does evolutionary plasticity evolve? *Evolution*, 50(3):1008–1023, 1996.

32. Andreas Wagner. Robustness and evolvability: a paradox resolved. *Proceedings of the Royal Society B: Biological Sciences*, 275(1630):91–100, 2008.

33. Andreas Wagner. *The origins of evolutionary innovations: a theory of transformative change in living systems.* OUP Oxford, 2011.

34. Andreas Wagner. Robustness and evolvability in living systems. 2013.

35. Günter P Wagner and Jianzhi Zhang. The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. *Nature Reviews Genetics*, 12(3):204–213, 2011.

36. P Wagner Gunter and L Altenberg. Complex adaptations and the evolution of evolvability. *Evolution*, 50(3), 1996.

37. Daniel B Weissman, Michael M Desai, Daniel S Fisher, and Marcus W Feldman. The rate at which asexual populations cross fitness valleys. *Theoretical population biology*, 75(4):286–300, 2009.

38. John J Welch and David Waxman. The nk model and population genetics. *Journal of theoretical biology*, 234(3):329–340, 2005.

39. Claus O Wilke. Robustness and evolvability in living systems. *BioScience*, 56(8):695–696, 2006.

40. Claus O Wilke and Christoph Adami. The biology of digital organisms. *TRENDS in ecology & evolution*, 17(11):528–532, 2002.

41. Claus O Wilke, Jia Lan Wang, Charles Ofria, Richard E Lenski, and Christoph Adami. Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature*, 412(6844):331–333, 2001.