scientific reports

OPEN

Path probability selection in nature and path integral

Chao Wang[1], Min-Lan Li[1,2] & Rui-Wu Wang[1✉]

Understanding of any biological evolutions, such as speciation, adaptation behavior and biodiversity pattern, is based on a fundamental concept of fitness, in which natural selection implies the improvement and progress of fitness in either direct/indirect benefit or genetic transmission to the next generation. However, this basic idea of biological evolution, which is mathematically described by Price equation or its related models, has not fully considered feedback effects from the environment or other generations. They lost the global dynamics of the evolutions consequently. Drawing on the idea of modern physics, we introduce the path integral by iterating the Price equation step by step to characterize the evolutionary path in which the stationary fitness is replaced by the path probability. The evolutionary selection therefore will depend on path probability instead of fitness advantage. In such a framework of the evolutionary path, the intermediate process of evolution is not always pointing to the fitness-maximizing equilibrium and multiple evolutionary paths could thus coexist without fitness advantage discrimination. This mechanism could potentially explain evolutionary strategies with the diversified fitness (e.g., coexistence of altruism and selfishness) and thus species diversity.

Evolution through natural selection is often understood to imply improvement and progress in fitness advantage which could be either direct/indirect fitness or genetic transmission to the next generation[1–5]. A trait, if it could increase the individual's fitness, will spread in the population and the average fitness of the population will therefore increase over time[2,4,6]. Such a metaphor of steady ascent on fitness landscape suggests some solid ground over which the population moves[7–9]. This scenario is described by Price equation mathematically, in which the character change is equal to the covariance between the character and its fitness (Fig. 1a)[3–5,10].

However, these traditional theories are resorted to the stationary fitness assumption and neglects some important evolutionary dynamics: although the environment selects the adaptations, these adaptations can shape the environment[9,11–14]. As Nowak and Sigmund metaphorized that selective advantage of a given tree height will also depend on the heights of neighboring trees, the fitness of the individuals will not just be affected by the phenotypic distributions of its own population, but also be greatly connected with the phenotype change of other species populations[9], as well as the interaction variation between ecological and biological factors[15–19]. When the fitness of a biological organism varies over either time or spatial scale, which might be resulted from both intrinsic genotype/phenotype uncertainty and environmental fluctuations, a uniform fitness throughout the whole evolution processes therefore might not be proper to describe evolutionary advantages of an organism[20]. The instantaneous fitness at a moment of the evolution, instead of uniform fitness as a stationary fitness, will be more appreciated and credible to describe the evolutionary process of a biological organism[21–23].

Instead of integrating frequency- or environment-dependence into population intrinsic fitness and concentrating the localized property near the equilibrium point (or steady state) (e.g., evolutionary game theory, adaptive dynamics, and population genetics)[20,24–27], here we state dynamics of the evolutionary history with the instantaneity of the fitness and focus on the evolutionary trajectories[22,28–31]. We put forward the evolutionary path probability to characterize the evolutionary trajectories. In this manner, each evolution step recurs taking the feedback effect among generations into account in place of the scenarios of a population evolving through the fitness advantage. In present work, we draw on a mathematical approach based on path-integral technique, which was introduced by Feynman for quantum physics and has been proved to be a very powerful tool in various areas of physics, both computationally and conceptually[32]. The path integral method, which was raised from Wiener[33] for Brownian motion, has found substantial success in applications in quantum mechanics and quantum field theory. In the 1940s Feynman[32] formulated a space–time approach to quantum mechanics based on the path integral introduced by himself. Recent years have witnessed a majority efforts to develop a similar path integral formulation for the theory of the stochastic process. And such a formulation can supply a calculational technique

[1]School of Ecology and Environment, Northwestern Polytechnical University, Xi'an 710072, Shaanxi, People's Republic of China. [2]School of Mathematics and Statistics, Northwestern Polytechnical University, Xi'an 710072, Shaanxi, People's Republic of China. ✉email: wangrw@nwpu.edu.cn
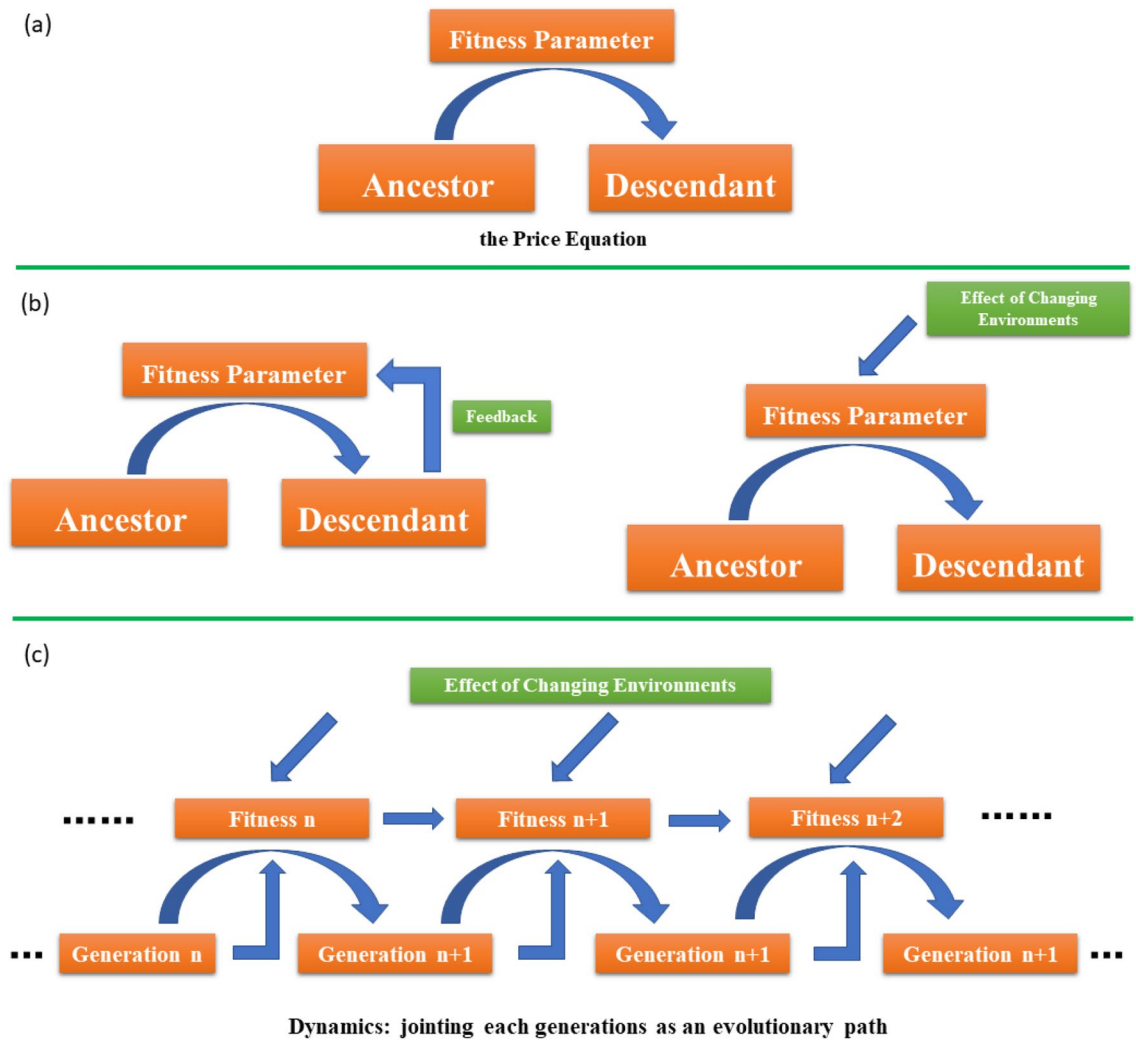
Figure 1. The Price equation and the evolutionary trajectory. (**a**) The Price equation describes evolutionary change between generations through the covariance of a character with its fitness. (**b**) The fitness of a generation might receive feedback originated from the change in quantity and character of the last generation or other more generations. The population distribution might also be affected by their living environments. (**c**) The whole evolutionary trajectory should be successive by jointing each generation step by step.

that may be fruitful in dealing with nonlinear systems. Although general evaluation of the path integral is impossible for nonlinear systems, its definition as the limit of a multidimensional integral suggests powerful numerical methods[34–36]. Recently, Joshua has performed attempting for the evolutionary problem of the gene frequency in the population genetics, showing how it is possible to use path integral methods to express the transition probability by means of a functional integral over paths of the evolutionary process for the Wright-Fisher process[37]. We also present a typical example on account of evolutionary game for approximating the evolutionary trajectory to make the conception more practical.

Mathematical background and the formulation of recursive evaluation

For sanity of our conceptions, we start from the simple Price's equation and then derivate the evolutionary path probability. In general, fitness is interpreted as the expected reproductive success of a trait[3,5]. The reproductive success is estimated in relation to a reference population: they are the linear combination of allelic frequencies that best explains phenotypic variation in the trait[3,38]. The Price's rule indicates that the evolutionary change in the average value of some traits, z, can be written as[3,5] (Fig. 1a)

$$\overline{\omega}\Delta\overline{z} = Cov(z,\omega) + E(\omega\Delta z), \tag{1}$$

where the fitness, ω, of a trait measures the number of its descendants at future time. The second term of the Price equation is sometimes added to account for non-Mendelian effects (e.g., meiotic drive) and random statistical deviations (e.g., genetic drift). To make the analysis more explicit, the second term is negligible without losing credibility in the following.
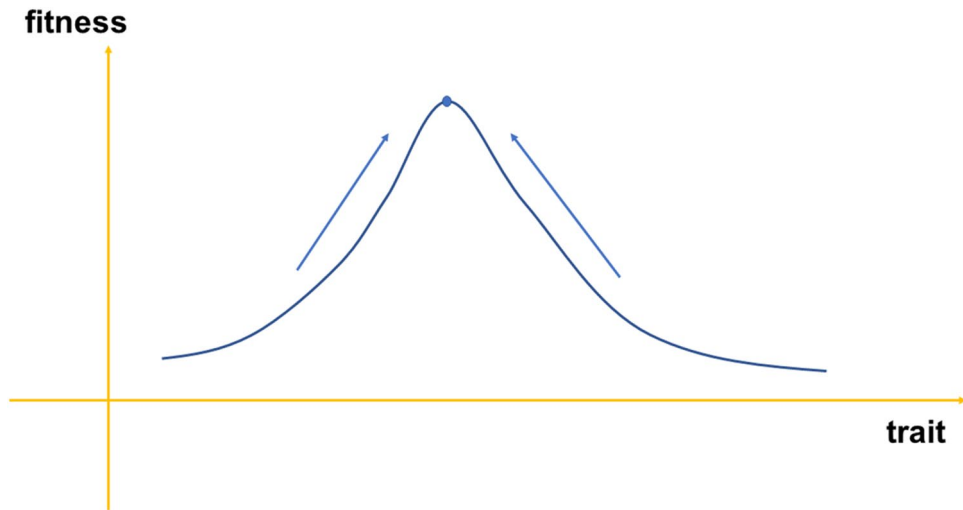
Figure 2. The evolutionary for the constant fitness distribution. In the traditional Price equation, where the fitness is constant, evolution always point to the equilibria. It is irrelevant to the traits. The evolution is only determined by the stationary fitness distributions and all the individual will evolve to position corresponding to the fitness maximum.

In the Price equation, the change in the average value of some property between two populations could be described. We consider a population (a set) in which each element is labeled by a label i. The label i may be used as any sort of property of trait z_i in the set, such as allele, genotype, phenotype, group of individuals, and so on. The frequency of elements with index i is q_i. For its descendant, the corresponding frequencies q_i' and the characters z_i' have the trait with the label i. We use Δq_i for frequency change associated with selection, and for property value change $\Delta z_i = 0$ as we neglect the mutation. The total change $\Delta \bar{z}$ is thus

$$
\begin{aligned}
\Delta \bar{z} = \bar{z}' - \bar{z} &= \sum q_i' z_i' - \sum q_i z_i \\
&= \sum q_i'(z_i' - z_i) + \sum q_i' z_i + \sum q_i z_i \\
&= \sum q_i' \Delta z_i + \sum \Delta q_i z_i.
\end{aligned}
\tag{2}
$$

By following the prior definition of fitness, we have

$$
\Delta q_i = q_i' - q_i = q_i \frac{\omega}{\bar{\omega}} - q_i = q_i \left(\frac{\omega}{\bar{\omega}} - 1 \right) = \frac{Cov(z, \omega)}{\bar{\omega}}.
\tag{3}
$$

This expression corresponds to the total change caused by changes in frequency. We call this the part caused by selection, because this is the part that arises directly from differential contribution by ancestors to the descendant population[5]. In this formular, the update information of fitness is not provided. In a stable fitness landscape or a short time, one can iterate the Price equation step by step with the fitness being regarded as a constant[38]. For the recursive evolutionary path, the fitness parameter ω of the evolutionary change between generations would receive feedback originated from the change in quantity and character of the last generation (and even other more generations), such as the evolution dynamic in the finite populations[9,20]. One should consider the nonlinear relationship (Fig. 1b)[27]. Also, the environmental impacts are included as a constant parameter for the fitness parameter in the traditional Price Equation, which would be liable to failure in long-term evolution (Fig. 1b). Instead of the fixed fitness, the fitness parameter dynamic also exists and the instantaneous fitness would offer a better way than the average fitness over time or stationary fitness[13]. We give the recurrence relation of the traits by substituting the instantaneous fitness $\omega^{(t)} = f[(q(t)] + \xi(t)$ instead of stationary fitness, where the first term comes from the feedback of population structure and $\xi(t)$ represents the impact of the environmental variation. The whole evolutionary process emerges from the recursive manner of the Price equation, which is a chain-like path (Fig. 1c).

In the traditional Price equation, where the fitness is constant, at length evolution reach the equilibrium point $Cov(z, \omega) = 0$. The evolution is determined by a stationary fitness distribution and the trait will always evolve to the fitness maximum point (shown in Fig. 2). At arbitrary time, the evolutionary direction is fixed toward the maximum point[2,4,5,29,30]. This evolutionary direction is irrelevant to the process. For a long evolutionary process and the fitness is frequency- dependent, the case changes. We assume that $q_i^{(t_0)}$ is the frequencies in the initial state, and $q_i^{(t_n)}$ is defined as the fraction of the descendant population at the generation n that is derived from members of the ancestral population that have the trait with the label i. If $i=1, 2$, $q_{1,2}^{(t_n)}$ is the fraction of the sequence descendants derived from entities with the phenotype $i=1, 2$ in the ancestors, therefore, i also
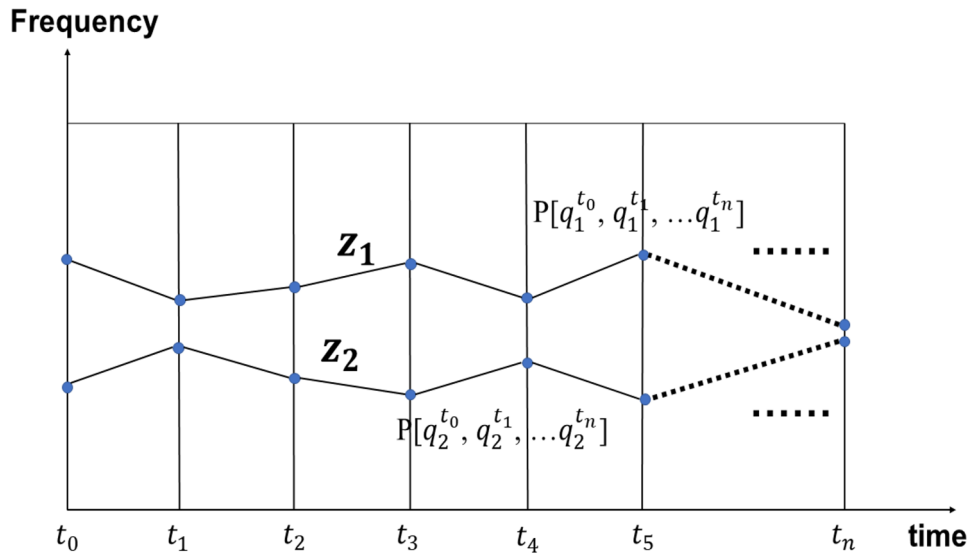
Frequency

Figure 3. The evolutionary trajectories. There are two evolutionary pathways from time t_0 to time t_n, say path z_1 and path z_2. Each evolutionary path appears with the possibility of $P[q_{1,2}^{(t)}]$ whose independent variable is a time dependent function [the evolutionary trajectories $q_{1,2}^{(t)}$]. For the discrete case, the distribution of the probability $P[q_{1,2}^{(t)}]$ can be written as $P_{path}[q_{1,2}^{(t_0)}, q_{1,2}^{(t_1)}, q_{1,2}^{(t_2)}, \ldots q_{1,2}^{(t_n)}]$. The path distribution is satisfied $q_1^{(t)} + q_2^{(t)}=1$ at each time point t.

represents a special evolution trajectory in the condition of neglecting random mutation. The fitness would rely on the frequency of the corresponding trait at last time with the recursion formula satisfied the Price equation:

$$q_i^{(t_n)} = q_i^{t_{n-1}}\left(\frac{w_i^{(t_n)}}{\overline{w}^{(t_n)}}\right). \tag{4}$$

$w_i^{(t_n)}(q_i^{(t_n)})$, above equation can be recombined as

$$q_i^{(t_n)} = q_i^{(t_0)}\prod_{k=0}^{n}\left(\frac{w_i^{(t_k)}}{\overline{w}^{t_k}}\right) \tag{5}$$

In this expression, the frequency of the individual in the i-th evolution trajectory, $q_i^{(t_0)}, q_i^{(t_1)}, q_i^{(t_2)}, \ldots q_i^{(t_n)}$ for discrete formula and $q_i^{(t_n)}$ for continuous case, receive effects from feedback on the fitness, including its own and others'. The evolution direction is no longer fixed and would depend on the choice of the path i, i.e., the path dependent selection.

Drawing on statistic physics, we introduce a conception of the probability distribution function (or rather functional) to identify the evolutionary path[32]. We sign this expanded generalized function distribution as P_{path} $[q_i^{(t_n)}]$, whose independent variable is a time dependent function [the evolution trajectories $q_i^{(t_n)}$][32]. Then, each trajectory has an evolutionary path probability, P_{path}, as shown in Fig. 3 (for simplicity $i = 1, 2$). This evolutionary-path probability $P_{path}[q_i^{(t_0)}, q_i^{(t_1)}, q_i^{(t_2)}, \ldots q_i^{(t_n)}]$ can be thought of as similar to the probability function for a number of dependent variables $q_i^{(t_0)}, q_i^{(t_1)}, q_i^{(t_2)}, \ldots,$ and $q_i^{(t_n)}$, which is an infinite dimensional function. This probability distribution function can be used to infer the preference for evolutionary trajectory in the evolutionary process. A comprehensible mathematical detail of the evolutionary path probability given by the path integral is shown in the "Method".

In fact, the population frequency q_i and the evolutionary path probability P_{path} are not consistent except in some extreme cases. At the time t_n, the individual adopts the phenotype i (Fig. 3) with the probability of $q_i^{(t_n)}$ ($\sum q_i^{(t_n)}=1$ at each time point t_n). By expanding to the distribution of pathway, one still could return to the trait distributions of $q_i^{(t_n)}$ at arbitrarily evaluation time point, $q_i^{(t_n)} = \sum_i P_{path}$. The mean value of the phenotypes at a given time point could be $\sum_i z_i(t)\sum_i P_{path}$ instead of $\sum_i z_i q_i$ in the Price equation, in which the path integral is instead by summation. For example, as shown in Fig. 4, the blue traits appear in $N+1$ generation with 50% while the probability of the path with the blue traits in $N+1$ generation is 6.25%. Only in one small step or a fitness stationary process, the phenotype strategy i is fixed, and $q_i^{(t_n)} = P_{path}[q_i^{(t_0)}, q_i^{(t_1)}, q_i^{(t_2)}, \ldots q_i^{(t_n)}]$.
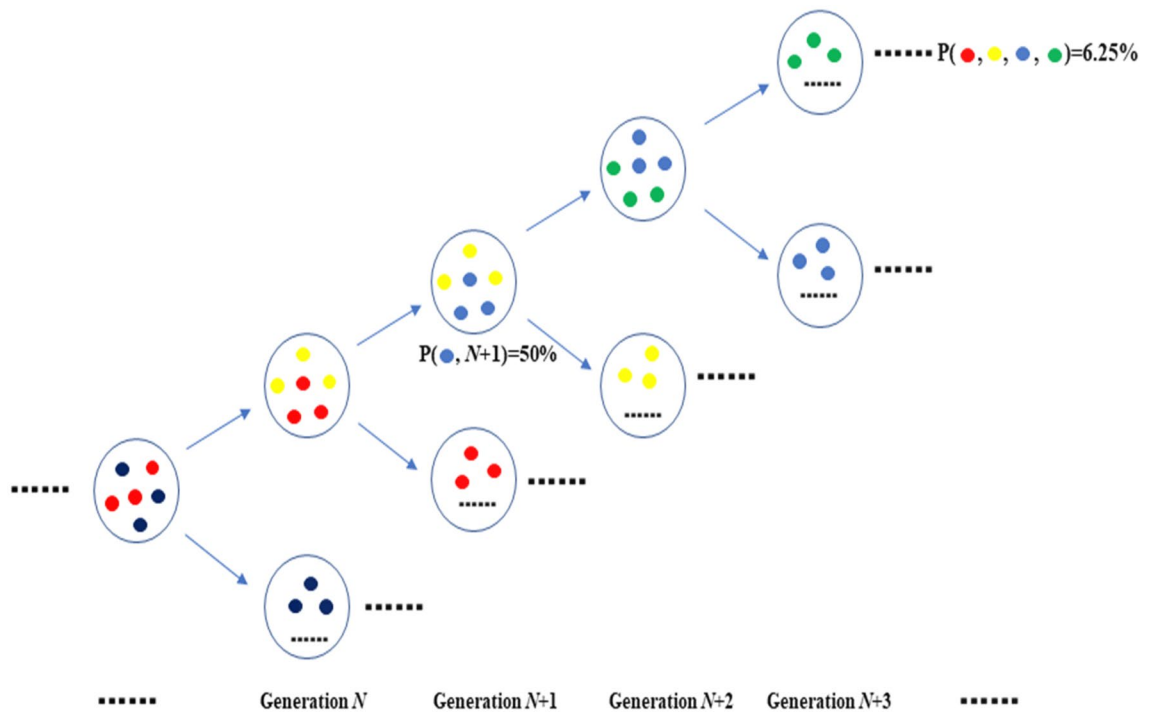
Figure 4. The probability distribution for strategies at a given time, P(x_i, t) and the probability of the path of evolutionary trajectories, P[x(t)]. The traits are marked by colors. At the generation $N + 1$, the frequency of the blue type is 50% in the population. At each time step, there are two phenotypes and each one is selected by the probability of 0.5. If we consider the four-generation case, the trajectory (the sequence "red, yellow, bule, green") with the frequency of the blue type being 50% in the generation $N + 1$ is 0.5 * 0.5 * 0.5 * 0.5 = 0.0625.

An example

We will take a simulation of evolutionary game to show the path probability, P_{path}, and give a practical execution below for the study of an evolutionary dynamics with the aid of the probability distributions of paths. We carry out the study in well-mixed finite populations with fixed size N. There are two states (the ancestral population), cooperation ($i = 1$) and defection ($i = 2$). The number of cooperatiors is denoted by m and updated by the Moran process[24]. Fitness is simply given from the expected payoff. We define the birth–death process given by the prisoner's dilemma with the payoff matrix ($c > a > d > b$):

$$\begin{array}{cc} & C\ D \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{array}.$$

Then, the expected payoff for $i = 1$ and 2 individuals are, respectively, given by

$$\begin{aligned} F_1 &= \frac{a(m-1) + b(N-m)}{N-1}, \\ F_2 &= \frac{cm + d(N-m-1)}{N-1}. \end{aligned} \tag{6}$$

In a finite population N, the fitness of the two strategies is frequency dependent. We also consider the selection intensity parameter h ($0 < h < 1$). The corresponding fitness at each step is then $w_{1,2} = 1 - h + hF_{1,2}$. This process has two "absorbing states" $m = 0$ and $m = N$: once the process has reached such a state, it will stay there forever. We investigate the trajectory from the initial state $m = N/2$ during 6000 simulation steps. In order to seek the pathways most likely to appear, we simulate 5000 evolutionary processes and analyzed the probability distribution of evolution trajectories by Monte Carlo. An evolutionary path can be described by a series of frequencies at successive history steps. In Fig. 5, we show the distribution of pathways under four levels of selection pressure.

From the simulation, one can see that the trajectory with maximum probability is located at the middle area of the curve. Although all the paths start from the same initial state and evolve to the same final state, they have different probabilities. If the selection pressure is not too strict (weak selection), the expected payoff $F_{1,2}$ is multi-value at each update step. The evolutionary path recured by the Price's rule is therefore not unique (In our simulation, the price's formula corresponds to the update rule). Then, a set of evolutionary paths could produce the same absorbing states. The pathways most likely appeared in the evolutionary process locate at the middle area and might not be derived by absorbing to the direction of equilibrium state $m = N$. That is exactly
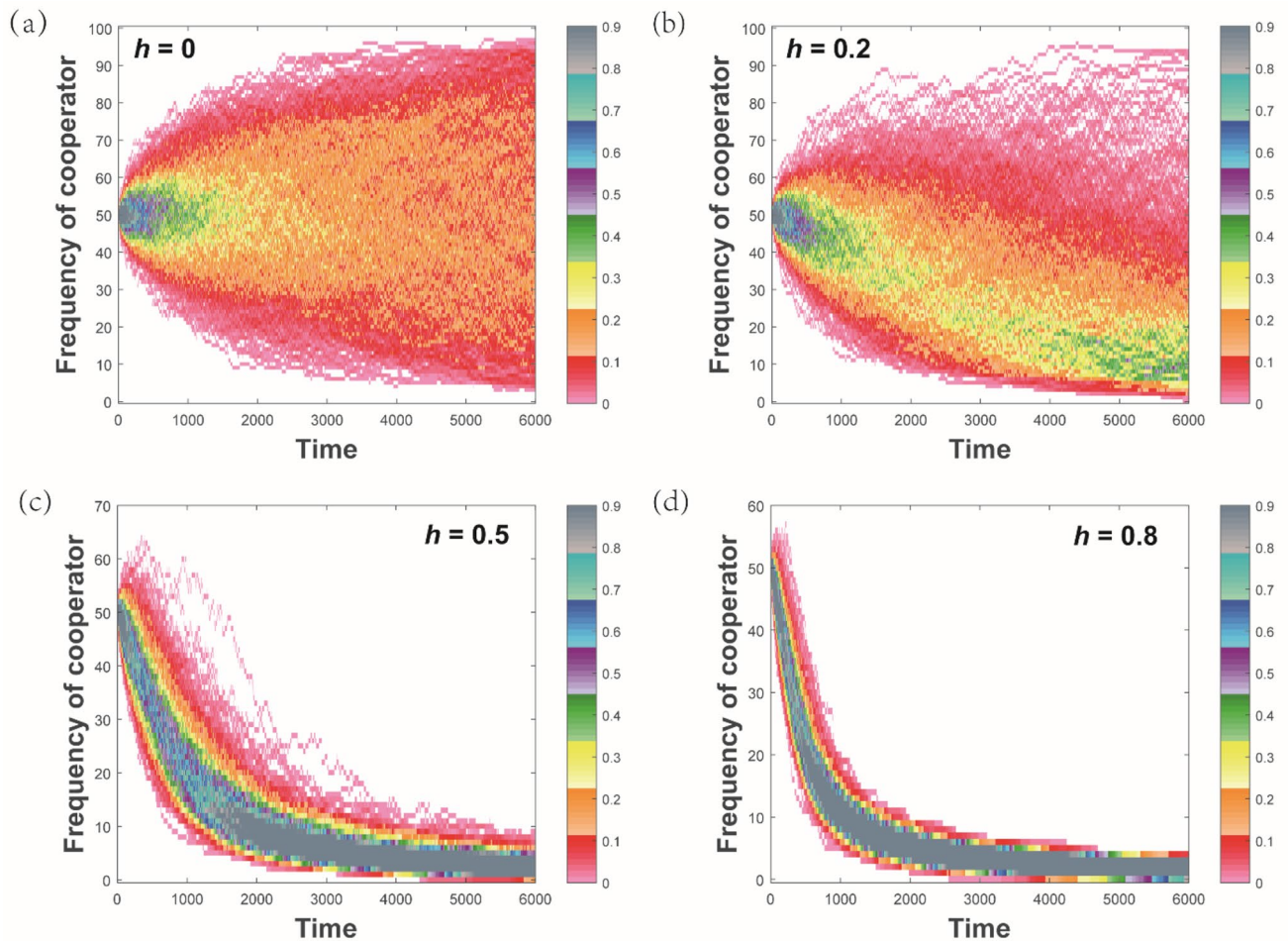
Figure 5. The distributions of the evaluation path probability under different selection pressure in the simulation of evolutionary game. (**a**–**d**) correspond to different selective pressure, $h = 0$, 0.2, 0.5, 0.8, respectively.

an exhibition of what the selection preference of the path-dependent selection is not completely consistent with the Price's description.

We consider different selection intensities, assuming that the species evolution is influenced by one or more factors (e.g., the interaction of different phenotypes, environment fluctuation), the parameter h being assigned ranging from 0 to 1. Small h means that there has low selection preference between cooperators and defectors at each step. For $h = 0$, all the evolutionary paths are produced with equal probability and selection is almost neutral. It is only constrained the initial value. With increasing selection pressure h, the distribution of paths that the strategy space cross would be more concentrated. For $h = 0.8$, only a few evolutionary paths are permitted and randomness decreases gradually. Whereas for an infinite great selection pressure, evaluation is allowed in a single path and the selection becomes directional.

Discussion

Evolutionary biology assumes that natural selection has already maximized fitness, in which we observe in nature is close to the trait with a fitness-maximizing equilibrium or the evolutionarily stable strategy[22]. Underlying this view is the assumption that the organism we observe is the final state of a long evolutionary process over many generations[22]. However, if there exist feedbacks between the adaptation and environments, such as the case of the finite population, the intermediate generations are a worthy addition to the evolution[28,31]. For the intermediate process of the speciation which we focus, evolutionary trajectory emphasizes that selection acts upon the probability distributions of paths rather than the phenotypes of the given equilibrium point[31]. Therefore, the evolutionary path probability parameter P_{path} is suitable for a sequence of evolutionary states and a more feasible patterns in the inquiry of the selection rule of the evolutionary paths.

In our framework, the evolutionary path contains the information of the junctions between the two repetitions which represent the change of fitness parameter. As mentioned by Wright, in a frequency-dependent selection the average fitness need not grow continually[39]. These junctions are thus informative and alter the fitness parameter[5,40]. Especially, evolutionary changing toward equilibria occurs slowly in the case of weak selection pressure (Fig. 5a,b). If the new ecological equilibria are generated before the population structure reaches ecological equilibria, the varying fitness makes the evolution not always stay at the maximum point. Some lower fitness traits

may thus appear in the evolutionary paths with lower probability. Unless we assume that population dynamics are typically considered operating at a much faster pace relative to evolutionary processes (ignoring effects of population dynamics on the fitness, such as infinite population size)[41], the junctions could not be dropped out.

In the meaning of the evolutionary paths, if the selection pressures are weak, the traits with low pay-offs might not be out-competed promptly in the intermediate process (Fig. 5). Thus, even an evolution starting from one initial state and falling into the same final state would allow for multiple intermediate paths. In fact, several experiments have been found including multiple different paths[21,39,42] and the multiple path evaluation will always occur in the substitution in individual base pairs of DNA[42]. In such a pattern, some possible evolutionary chains might develop along with lower reproductive rates even if natural selection favors the trait with increasing fitness at each evolutionary step. When the evolution is multiple-pathway, it means that the expected fitness is the same but the same intermediate state is not exactly equal, i.e., sympatric species, may exist.

Since the distribution of evolutionary path probability P_{path} is a multi-dimensional function taking the frequency $q_i^{(t)}$ as the dependent variables at the whole evolutionary step, it will extend the conception of species to speciation. This description makes the essence of any evolution to a process where the populations will not only keep their divergence upon contact but also be able to continue to diverge[43,44]. It contains a unification of the statics and dynamic properties in an evolutionary process. We should stress that the stationary population distribution $q_i^{(t_n)}$ at a given time is equivalent to the evolutionary path probability P_{path} only in one small step or a fitness stationary case. Specifically, if the selection pressure is small, all evolutionary paths occur in a nearly identical probability (Fig. 5a). With the increase of selection pressure, the distribution of the paths is more concentrated and the direction of the evolutionary path shows up gradually. If the selection pressure is strong enough, evaluation is allowed in a single path and the becomes directional, i.e., the traditional Price equation[13].

In conclusion, we focus on the selection rule based on the path-depended. Stating from the path-depended selection, we found that the evolutionary path expands several missing but meaningful phenomena in the fitness-based theory. Drawing on statistic physics, we suggest that the conception of the probability distribution function (or rather functional) could be introduced to identify the evolutionary path. Considering the evolutionary path, some individuals with lower fitness could coexist with others due to they get a large evolutionary path probability P_{path} if selective pressure of the natural selection is not too strong. The distinction selective mechanism of the evolutionary path probability in the whole evolutionary process might seed a potential to explain the diversified fitness evolutionary strategies (e.g., coexistence of altruism and selfishness) and thus species diversity.

Method

We could give a mathematical formalism based on the probability function of evolution trajectories. In general, the growth value can be written as

$$f(x, t) + g(x, t)\xi(t), \tag{7}$$

where $f(x, t)$ and $g(x, t)$ are smooth functions, $f(x, t)$ involved fitness (received feedback from other generations), and $\xi(t)$ is the term for environmental fluctuations. We consider an evolution process.

$$\dot{x} = f(x, t) + g(x, t)\xi(t) \tag{8}$$

If the phenotype began with x_0 at t_0, taking account of the path integral, we obtain probability of evolutionary when the final phenotype is x at t

$$P(x, t; x_0, t_0) = \int D[x(t)]\exp\left[-\frac{1}{h}\int_{t_0}^{t} d\tau L[x(t), \dot{x}(t)]\right],$$

with $L[x(t), \dot{x}(t)]$ being the so called Lagrangian in Lagrangian mechanics analysis. $L[x(t), \dot{x}(t)]$ is another form of the Eq. (8). The probability densities of each trajectory are, following Ref.[45],

$$P[x(t)] = D[x(t)]\exp\left[-\frac{1}{h}\int_{t_0}^{t} d\tau L[x(t), \dot{x}(t)]\right].$$

It seems that this is a complex formula, but still gives a general mathematic formula and practical execution for study of the distribution of an evolutionary trajectory. In fact, the evolutionary significances of the path probability would become clearly in a specific system as our contrived example of evolutionary game simulation.

Data availability

This work is pure theoretical research and there is no experimental data. All data generated or analyzed during this study are included in this published article. The programs for the current study are available from the corresponding author on reasonable request.

References

1. Darwin, C. R. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (John Murray, 1859).
2. Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon Press, 1930).
3. Price, G. R. Selection and covariance. *Nature* **227**(5257), 520–521 (1970).

4. Price, G. R. Extension of covariance selection mathematics. *Ann. Hum. Genet.* **35**, 485–490 (1972).
5. Frank, S. A. Natural selection. IV. The price equation. *J. Evol. Biol.* **25**(6), 1002–1019 (2012).
6. Frank, S. A. *Foundations of Social Evolution* (Princeton University Press, 2019).
7. Wright S. The roles of mutation, inbreeding, crossbreeding and selection in evolution. In *Sixth International Congress on Genetics* Vol. 1, No. 6, 356–366 (1932).
8. Eigen, M., Mccaskill, J. & Schuster, P. Molecular quasi-species. *J. Phys. Chem.* **92**(24), 6881–6891 (1988).
9. Nowak, M. A. & Sigmund, K. Evolutionary dynamics of biological games. *Science* **303**(5659), 793–799 (2004).
10. Lehtonen, J., Okasha, S. & Helanterä, H. Fifty years of the price equation. *Philos. Trans. R. Soc. B Biol. Sci.* **375**(1797), 20190350 (2020).
11. Levin, S. A., Grenfell, B., Hastings, A. & Perelson, A. S. Mathematical and computational challenges in population biology and ecosystems science. *Science* **275**(5298), 334–343 (1997).
12. Levin, B. R., Lipsitch, M. & Bonhoeffer, S. Population biology, evolution, and infectious disease: Convergence and synthesis. *Science* **283**(5403), 806–809 (1999).
13. Rice, S. H. A stochastic version of the Price equation reveals the interplay of deterministic and stochastic processes in evolution. *BMC Evol. Biol.* **8**(1), 262 (2008).
14. Nowak, M. A. & Highfield, R. *Super Cooperators: Altruism, Evolution, and Why We Need Each Other to Succeed* (Free Press, 2011).
15. Barton, N. H. & Turelli, M. Adaptive landscapes, genetic distance and the evolution of quantitative characters. *Genet. Res.* **49**(2), 157–173 (1987).
16. Hairston, N. G. Jr. *et al.* Rapid evolution revealed by dormant eggs. *Nature* **401**(6752), 446 (1999).
17. Hairston, N. G. Jr., Ellner, S. P., Geber, M. A., Yoshida, T. & Fox, J. A. Rapid evolution and the convergence of ecological and evolutionary time. *Ecol. Lett.* **8**(10), 1114–1127 (2005).
18. Schoener, T. W. The newest synthesis: Understanding the interplay of evolutionary and ecological dynamics. *Science* **331**(6016), 426–429 (2011).
19. Hendry, A. P. *Eco-evolutionary Dynamics* (Princeton University Press, 2017).
20. Taylor, C., Fudenberg, D., Sasaki, A. & Nowak, M. A. Evolutionary game dynamics in finite populations. *Bull. Math. Biol.* **66**(6), 1621–1644 (2004).
21. Poelwijk, F., Kiviet, D., Weinreich, D. & Tans, S. Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* **445**(7126), 383–386 (2007).
22. Traulsen, A., Iwasa, Y. & Nowak, M. A. The fastest evolutionary trajectory. *J. Theor. Biol.* **249**(3), 617–623 (2007).
23. Gokhale, C. S., Iwasa, Y., Nowak, M. A. & Traulsen, A. The pace of evolution across fitness valleys. *J. Theor. Biol.* **259**(3), 613–620 (2009).
24. Moran, P. A. P. Random processes in genetics. *Math. Proc. Camb. Philos. Soc.* **54**(1), 60–71 (1958).
25. Nowak, M. A. & May, R. M. Evolutionary games and spatial chaos. *Nature* **359**, 826–829 (1992).
26. Dieckmann, U. & Doebeli, M. A. On the origin of species by sympatric speciation. *Nature* **400**(6742), 354–357 (1999).
27. Hilbe, C., Šimsa, Š, Chatterjee, K. & Nowak, M. A. Evolution of cooperation in stochastic games. *Nature* **559**(7713), 246–249 (2018).
28. Gould, S. J. The evolution of life on the earth. *Sci. Am.* **271**(4), 84–91 (1994).
29. Mustonen, V. & Lässig, M. Fitness flux and ubiquity of adaptive evolution. *Proc. Natl. Acad. Sci.* **107**(9), 4248–4253 (2010).
30. Steinberg, B. & Ostermeier, M. Environmental changes bridge evolutionary valleys. *Sci. Adv.* **2**(1), e1500921 (2016).
31. Li, M. L., Wang, C. & Wang, R. W. Path-dependent speciation in the process of evolution. *Biodivers. Sci.* **29**(3), 409–418 (2021).
32. Feynman, R. P. Space-time approach to non-relativistic quantum mechanics. *Rev. Mod. Phys.* **20**(2), 367–387 (1948).
33. Wiener, N. The average of an analytic functional and the brownian movement. *Proc. Natl. Acad. Sci. USA.* **7**, 10 (1921).
34. Wehner, M. F. Numerical evaluation of path integral solutions to fokker-planck equations with application to void formation. *Rev. Mod. Phys.* **27**, 5 (1983).
35. Wehner, M. F. & Wolfer, W. G. Numerical evaluation of path-integral solutions to Fokker-Planck equations. II. Restricted stochastic processes. *Phys. Rev. A* **28**, 5 (1983).
36. Wehner, M. F. & Wolfer, W. G. Numerical evaluation of path-integral solutions to Fokker-Planck equations. III. Time and functionally dependent coefficients. *Phys. Rev. A* **35**, 5 (1987).
37. Schraiber, J. G. A path integral formulation of the Wright-Fisher process with genic selection. *Theor. Popul. Biol.* **92**, 30–35 (2014).
38. Queller, D. C. A general model for kin selection. *Evolution* **46**(2), 376–380 (1992).
39. Weinreich, D. M. The rank ordering of genotypic fitness values predicts genetic constraint on natural selection on landscapes lacking sign epistasis. *Genetics* **171**(3), 1397–1405 (2005).
40. Frank, S. A. George Price's contributions to evolutionary genetics. *J. Theor. Biol.* **175**, 373–388 (1995).
41. Pasztor, L., Botta-Dukát, Z., Magyar, G., Czárán, T. & Meszéna, G. *Theory-Based Ecology: A Darwinian Approach* (Oxford University Press, 2016).
42. Weinreich, D. M., Delaney, N. F., Depristo, M. A. & Hartl, D. L. Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* **312**(5770), 111–114 (2006).
43. He, Z. *et al.* Speciation with gene flow via cycles of isolation and migration: Insights from multiple mangrove taxa. *Natl. Sci. Rev.* **6**(2), 275–288 (2019).
44. Wang, X., He, Z., Shi, S. & Wu, C. I. Genes and speciation: Is it time to abandon the biological species concept?. *Natl. Sci. Rev.* **7**(8), 1387–1397 (2020).
45. Feynman, R. P., Hibbs, A. R. & Styer, D. F. *Quantum mechanics and path integrals* (Courier Corporation, 2010).

Author contributions

C.W. and R.W.W. wrote the main manuscript text and C.W. prepared Figs. 1, 2, 3, 4. M.L.L. simulated and make the Fig. 5. All authors reviewed the manuscript.C.W. and M.L.L. contributed equally to this work.

Funding

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to R.-W.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.