

# Evolution of cooperation by phenotypic similarity

Tibor Antal<sup>a</sup>, Hisashi Ohtsuki<sup>b,c</sup>, John Wakeley<sup>d</sup>, Peter D. Taylor<sup>e</sup>, and Martin A. Nowak<sup>a,d,1</sup>

<sup>a</sup>Program for Evolutionary Dynamics and Department of Mathematics, Harvard University, Cambridge, MA 02138; <sup>b</sup>Department of Value and Decision Science, Tokyo Institute of Technology, Tokyo 152-8552, Japan; <sup>c</sup>Precursory Research for Embryonic Science and Technology, Japan Science and Technology Agency, Saitama 332-0012, Japan; <sup>d</sup>Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138; and <sup>e</sup>Department of Mathematics and Statistics, Queen's University, Kingston, ON, Canada K7L 3N6

Communicated by Simon A. Levin, Princeton University, Princeton, NJ, March 10, 2009 (received for review June 15, 2008)

**The emergence of cooperation in populations of selfish individuals is a fascinating topic that has inspired much work in theoretical biology. Here, we study the evolution of cooperation in a model where individuals are characterized by phenotypic properties that are visible to others. The population is well mixed in the sense that everyone is equally likely to interact with everyone else, but the behavioral strategies can depend on distance in phenotype space. We study the interaction of cooperators and defectors. In our model, cooperators cooperate with those who are similar and defect otherwise. Defectors always defect. Individuals mutate to nearby phenotypes, which generates a random walk of the population in phenotype space. Our analysis brings together ideas from coalescence theory and evolutionary game dynamics. We obtain a precise condition for natural selection to favor cooperators over defectors. Cooperation is favored when the phenotypic mutation rate is large and the strategy mutation rate is small. In the optimal case for cooperators, in a one-dimensional phenotype space and for large population size, the critical benefit-to-cost ratio is given by  $b/c = 1 + 2/\sqrt{3}$ . We also derive the fundamental condition for any two-strategy symmetric game and consider high-dimensional phenotype spaces.**

coalescent theory | evolutionary dynamics | evolutionary game theory | mathematical biology | stochastic process

**E**volutionary game theory is the study of frequency-dependent selection (1–8). Fitness values depend on the relative abundance, or frequency, of various strategies in the population, for example, the frequency of cooperators and defectors. Evolutionary game theory has been applied to understand the evolution of cooperative interactions in viruses, bacteria, plants, animals, and humans (9–13). The classical approach to evolutionary game dynamics assumes well-mixed populations, where every individual is equally likely to interact with every other individual (4). Recent advances include the extension to populations that are structured by geography or other factors (14–25).

The term “greenbeard effect” was coined in sociobiology to describe the result of the following thought experiment (26, 27). What evolutionary dynamics will occur if a single gene is responsible for both a phenotypic signal (“a green beard”) and a behavioral response (for example, altruistic behavior toward individuals with like phenotypes)? Later, the term “armpit effect” was introduced (28) to refer to a self-referent phenotype that is used in identifying kin (29–31).

Both of these concepts are now seen as cases of “tag-based cooperation,” in which a generic system of phenotypic tags is used to indicate similarity or difference, and the evolutionary dynamics of cooperation are studied in the context of these tags. A first approach, based on computer simulations, assumed a well-mixed population, a continuum of tags, and an evolving threshold distance for cooperation (32). More recent models use numerical and analytic methods and often combine tags with viscous population structure (33–37). A general finding of these articles is that it is difficult to obtain cooperation in tag-based models for well-mixed populations, indicating that some spatial structure is needed (14).

Inspired by work on tag-based cooperation (32–34, 38) and building on a previous approach (39), we study evolutionary game dynamics in a model where the behavior depends on phenotypic distance (40, 41). As a particular example we explore the evolution of cooperation (42, 43). Studies of different organisms, including humans, support the idea that cooperation is more likely among similar individuals (31, 44–49). Our model applies to situations where individuals tend to like those who have similar attitudes and beliefs. We introduce a natural model in which individuals mutate to adjacent phenotypes in a possibly multidimensional phenotype space. We study one and infinitely many dimensions in detail. We develop a theory for general evolutionary games, not just the evolution of cooperation. Spatial structure is not needed for cooperation to be favored in our model. Moreover, in contrast to previous work (39), we develop an analytic machinery for describing heterogeneous populations in phenotype space.

Consider a population of asexual haploid individuals, with a population size  $N$  that is constant over time. Each individual is characterized by a phenotype, given by an integer  $i$  that can take any value from minus to plus infinity. Thus, this phenotype space is a one-dimensional and unbounded lattice. Individuals inherit the phenotype of their parent subject to some small variation. If the parent's phenotype is  $i$ , then the offspring has phenotype  $i - 1$ ,  $i$ , or  $i + 1$  with probabilities  $v$ ,  $1 - 2v$ , and  $v$ , respectively. The parameter  $v$  can vary between 0 and 1/2.

Let us consider a Wright–Fisher process. In each generation, all individuals produce the same large number of offspring. The next generation of  $N$  individuals is sampled from this pool of offspring. To introduce some fundamental concepts and quantities, we first study the model without any selection. No evolutionary game is yet being played, and there is only neutral drift in phenotype space. The entire population performs a random walk with a diffusion coefficient  $v$ , and by this process will tend to disperse over the lattice. In opposition to this, all of the individuals in the population will be, to some degree, related due to reproduction in a finite population. Thus, while occasionally the population may break up into two or more clusters, typically there is only a single cluster (50, 51). The standard deviation of the distribution in phenotype space, which is a measure for the width of the cluster, is  $\sqrt{2Nv}$ .

Next, we superimpose the neutral drift of two types: the strategies  $A$  and  $B$  (Fig. 1). Still for the moment, assuming no fitness differences, we have reproduction subject to mutation between  $A$  and  $B$ . Specifically, with probability  $u$  the offspring adopts a random strategy. The mutation–reproduction process defines a stationary distribution (52). If  $u$  is very small relative to  $N$ , the population tends to be either all- $A$  or all- $B$ . If  $u$  is large, the population tends toward one-half  $A$  and one-half  $B$ . Fig. 2 illustrates the random walk in phenotype space of the population composed of the two types  $A$  and  $B$ .

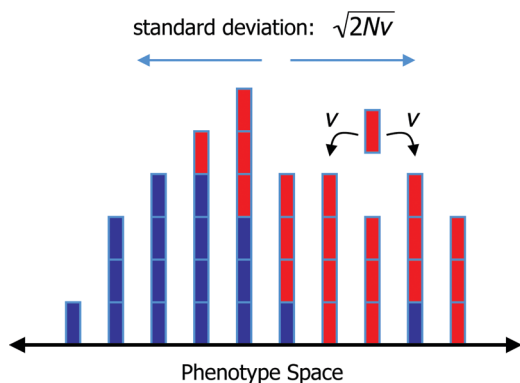
Author contributions: T.A., H.O., J.W., P.D.T., and M.A.N. performed research; and T.A., H.O., J.W., P.D.T., and M.A.N. wrote the paper.

The authors declare no conflict of interest.

See Commentary on page 8405.

<sup>1</sup>To whom correspondence should be addressed. E-mail: martin\_nowak@harvard.edu.

This article contains supporting information online at [www.pnas.org/cgi/content/full/0902528106/DCSupplemental](http://www.pnas.org/cgi/content/full/0902528106/DCSupplemental).



**Fig. 1.** The basic geometry of evolution in phenotype space. There are two types of individuals (red and blue), which can refer to arbitrary traits or different strategies in an evolutionary game. Individuals inherit the strategy of their parent subject to a small mutation rate  $u$ . Moreover, each individual has a phenotype. Here, we consider a discrete one-dimensional phenotype space. An individual of phenotype  $i$  produces offspring of phenotype  $i - 1$ ,  $i$ , or  $i + 1$  with probabilities  $v$ ,  $1 - 2v$ , and  $v$ , respectively. The total population (of size  $N$ ) performs a random walk in phenotype space with diffusion coefficient  $v$ . Sometimes the cluster breaks into two or more pieces, but typically only one of them survives. If evolutionary updating occurs according to a Wright-Fisher process then the distribution of individuals in phenotype space has a standard deviation of  $\sqrt{2Nv}$ . For the Moran process, the standard deviation is reduced to  $\sqrt{Nv}$ .

By using coalescence theory (53, 54) many interesting and relevant properties of the distributions of both the strategies and phenotypic tags can be calculated. For example, the probability that two randomly chosen individuals have the same phenotype is  $z = 1/\sqrt{8Nv}$ . The probability that two randomly chosen individuals have the same strategy and the same phenotype is  $g = z(1 - Nu/2)$ . The probability that two individuals have the same strategy and a third individual has the same phenotype as the second is  $h = z[1 - Nu(2 + \sqrt{3})/4]$ . These results hold for large population size  $N$  and small mutation rate  $u$ ; more precisely, we assume large  $Nv$  and small  $Nu$ . The relevance of  $z$ ,  $g$ , and  $h$  will become clear below. The expressions for  $z$ ,  $g$ , and  $h$  are derived for general  $Nv$  and  $Nu$  in supporting information (SI) Appendix, where they appear as Eqs. 10, 19, and 24, respectively.

We can now use these insights to study game dynamics. We investigate the competition of cooperators,  $C$ , and defectors,  $D$ . Cooperators play a conditional strategy: they cooperate with all individuals who are close enough in phenotype space and defect otherwise. The notion of being close enough is modeled by a lattice structure. In particular, a cooperator with phenotype  $i$  cooperates only with other individuals of phenotype  $i$ . Defectors, in contrast, play an unconditional strategy: they always defect. Cooperation means paying a cost,  $c$ , for the other individual to receive a benefit  $b$ . The larger the total payoff of an individual, interacting equally with every member of the population, the larger the number of offspring it will produce on average. We want to calculate the critical benefit-to-cost ratio,  $b/c$ , that allows the game in phenotype space to favor the evolution of cooperation.

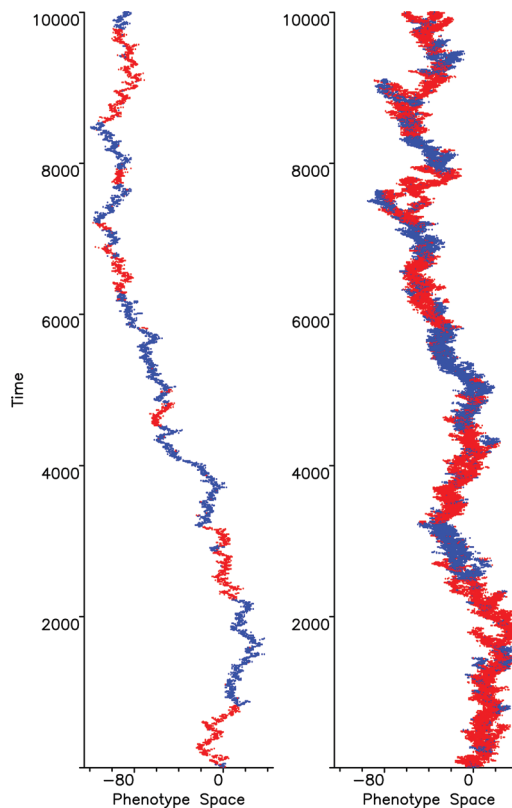
A configuration of the population is specified by  $m_i$  and  $n_i$ , which are the number of cooperators with phenotype  $i$  and the total number of individuals with phenotype  $i$ , respectively. The total payoff of all cooperators is  $F_C = \sum_i m_i (bm_i - cn_i)$ . The total payoff of all defectors is  $F_D = \sum_i (n_i - m_i)bm_i$ . There are  $\sum_i m_i$  cooperators and  $N - \sum_i m_i$  defectors. The average payoff for a cooperator is  $f_C = F_C / \sum_i m_i$ . The average payoff for a defector is  $f_D = F_D / (N - \sum_i m_i)$ . Cooperators have a higher fitness than defectors if  $f_C > f_D$ , which leads to  $\sum_i m_i (bm_i - cn_i) > \sum_i m_i \sum_j m_j n_j (b - c) / N$ . Averaging these quantities over every possible configuration of the population, weighted by their stationary probability under neutrality, we obtain the fundamental condition

$$b \left\langle \sum_i m_i^2 \right\rangle - c \left\langle \sum_i m_i n_i \right\rangle > (b - c) \left\langle \sum_{ij} m_i m_j n_j \right\rangle / N. \quad [1]$$

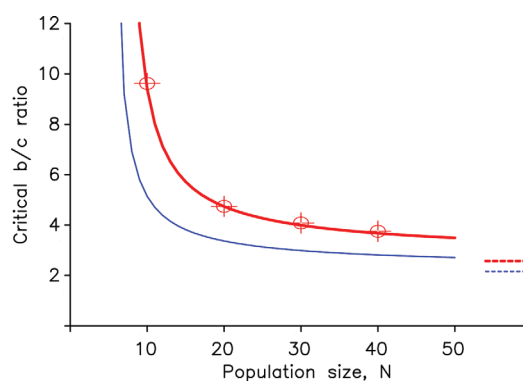
Under this condition cooperators are more abundant than defectors in the mutation-selection process. The above argument and our results are valid in the weak selection limit. A precise derivation of this inequality is presented in SI Appendix. Correlation terms similar to the ones above sometimes arise in studies of social behavior and population dynamics (26, 55). The first two terms in inequality (Eq. 1) are pairwise correlations, while the third is notably a triplet correlation. Note that the argument leading to inequality (Eq. 1) includes self-interaction, but that the effect of this becomes negligible when  $N$  is large.

When the population size is large, the averages in inequality (Eq. 1) are proportional to the probabilities  $g$ ,  $z$ , and  $h$  respectively, which we introduced earlier. Consequently, inequality (Eq. 1) can be written as  $bg - cz > (b - c)h$ . Using the values of  $z$ ,  $g$ , and  $h$  given above we obtain

$$b/c > 1 + \frac{2}{\sqrt{3}}, \quad [2]$$



**Fig. 2.** Random walks in phenotype space. Shown are two computer simulations of a Wright-Fisher process in a one-dimensional discrete phenotype space. The phenotypic mutation rate is  $v = 0.25$ . The colors, red and blue, refer to arbitrary traits, because no game is yet being played. All individuals have the same fitness. The population size is (Left)  $N = 10$ , and (Right)  $N = 100$ . The strategy mutation probability (between red and blue) is  $u = 0.004$ . Therefore, a given color dominates on average for  $2/u = 500$  generations (since new mutations arrive at rate  $Nu/2$  and fixate with probability  $1/N$ ). The standard deviation of the distribution in phenotype space is  $\sqrt{2Nv}$ . Approximately 95% of all individuals are within 4 standard deviations. Often the population fragments into two or several pieces, but only one branch survives in the long run. We use the statistics of these neutral “phenotypic space walks” for calculating the fundamental conditions of evolutionary games in the limit of weak selection.



**Fig. 3.** Excellent agreement between numerical simulations and analytic calculations. We show the critical benefit-to-cost ratio that is needed for cooperators to be more abundant than defectors in the stationary distribution. We have used a Wright–Fisher process with a phenotypic mutation rate  $\nu = 1/2$  and a strategy mutation probability  $u = 1/(2N)$ . The red line indicates the result of our analytic calculation. For these parameter values the asymptotic limit for large  $N$  is  $b/c = (1 + 12\sqrt{2})/7 \approx 2.5672$ . The red dots indicate the result of numerical simulations. The gray line illustrates the critical  $b/c$ -ratio for  $u \rightarrow 0$  with the asymptotic limit  $b/c = 1 + 2/\sqrt{3} \approx 2.1547$ .

which is approximately 2.16. If the benefit-to-cost ratio exceeds this number, then cooperators are more abundant than defectors in the mutation-selection process. The success of cooperators results from the balance of movement and clustering in phenotype space. Inequality (Eq. 2) represents the condition for cooperators to be more abundant than defectors in a large population when the strategy mutation rate  $u$  is small ( $Nu \ll 1$ ) and the phenotypic mutation rate  $\nu$  is large ( $N\nu \gg 1$ ). In *SI Appendix*, we derive conditions for any population size and mutation rates. Fig. 3 shows the excellent agreement between numerical simulations and analytical calculations. In general, we find that both lowering strategy mutations and increasing phenotypic mutations favor cooperators.

We can expand our analysis to study any  $2 \times 2$  game, not only the interaction between cooperators and defectors. Consider two strategies  $A$  and  $B$  and the general payoff matrix

$$\begin{array}{cc} & \begin{array}{c} A \\ B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{array} \quad [3]$$

The payoffs for  $A$  versus  $A$ ,  $A$  versus  $B$ ,  $B$  versus  $A$ , and  $B$  versus  $B$  are given by  $R$ ,  $S$ ,  $T$ ,  $P$ , respectively.  $A$  players use strategy  $A$  against other individuals with the same phenotype, otherwise they use  $B$ .  $B$  players always use strategy  $B$ . For the game in a one-dimensional phenotype space and large population size we find that  $A$  is more abundant than  $B$  if

$$(R - P)(1 + \sqrt{3}) > T - S. \quad [4]$$

For the derivation see Section 5.2 in the *SI Appendix*. This formula can be used for evaluating any two-strategy symmetric game in a one-dimensional phenotype space. In the *SI Appendix*, we discuss the snow-drift game and the stag-hunt game as particular examples.

We can also study higher-dimensional phenotype spaces. In general, for higher dimensions, it is easier for cooperators to overcome defectors. The intuitive reason is that in higher dimensions phenotypic identity also implies strategic identity. In Section 5.3 of the *SI Appendix*, we show that, in the limit of infinitely many dimensions, and under the same assumptions that produced conditions 2 and 4, the crucial benefit-to-cost ratio in the Prisoner's Dilemma converges to  $b/c > 1$ . For general games, the equivalent result of condition 4 becomes  $R > P$ , which means the evolutionary process always chooses the strategy with the higher payoff against itself. Our basic approach can also be adapted to continuous, rather than discrete, phenotype spaces. In this case, no two individuals have exactly the same phenotype, but the conditional behavioral strategy is triggered by sufficient phenotypic similarity.

In summary, we have developed a model for the evolution of cooperation based on phenotypic similarity. Our approach builds on previous ideas of tag-based cooperation, but in contrast to earlier work (33–37), we do not need spatial population dynamics to obtain an advantage for cooperators. We derive a completely analytic theory that provides general insights. We find that the abundance of cooperators in the mutation-selection equilibrium is an increasing function of the phenotypic mutation rate and a decreasing function of the strategic mutation rate. These observations agree with the basic intuition that higher phenotypic mutation rates reduce the interactions between cooperators and defectors, whereas higher strategic mutation rates destabilize clusters of cooperators by allowing frequent invasion of newly mutated defectors. Therefore, cooperation is more likely to evolve if the strategy mutation rate is small and if the phenotypic mutation rate is large. In a genetic model this assumption may be fulfilled if the strategy is encoded by one or a few genes, whereas the phenotype is encoded by many genes. Also in a cultural model, it can be the case that the phenotypic mutation rates are higher than the strategic mutation rates; for example, people might find it easier to modify their superficial appearance than their fundamental behaviors. Furthermore, we show how the correlations between strategies and phenotypes can be obtained from neutral coalescence theory under the assumption that selection is weak (54, 56). Our theory can be applied to study any evolutionary game in the context of conditional behavior that is based on phenotypic similarity or difference.

**ACKNOWLEDGMENTS.** This work was supported by the John Templeton Foundation, the National Science Foundation/National Institutes of Health (R01GM078986) joint program in mathematical biology, the Bill and Melinda Gates Foundation (Grand Challenges Grant 37874), the Japan Society for the Promotion of Science, and J. Epstein.

- Smith JM, Price GR (1973) The logic of animal conflict. *Nature* 246:15–18.
- Taylor PD, Jonker L (1978) Evolutionarily stable strategies and game dynamics. *Math Biosci* 40:145–156.
- Smith JM (1982) *Evolution and the Theory of Games* (Cambridge Univ Press, Cambridge, UK).
- Hofbauer J, Sigmund K (1998) *Evolutionary Games and Population Dynamics* (Cambridge Univ Press, Cambridge, UK).
- Cressman R (2003) *Evolutionary Dynamics and Extensive Form Games* (MIT Press, Cambridge, MA).
- Vincent TL, Brown JS (2005) *Evolutionary Game Theory, Natural Selection, and Darwinian Dynamics* (Cambridge Univ Press, Cambridge, UK).
- Nowak MA, Sigmund K (2004) Evolutionary dynamics of biological games. *Science* 303:793–799.
- May RM (1973) *Stability and Complexity in Model Ecosystems* (Princeton Univ Press, Princeton, NJ).
- Parker GA (1974) Assessment strategy and evolution of fighting behavior. *J Theor Biol* 47:223–243.
- Colman AM (1995) *Game Theory and Its Applications in the Social and Biological Sciences* (Butterworth–Heinemann, Oxford).
- Sinervo B, Lively CM (1996) The rock-paper-scissors game and the evolution of alternative male strategies. *Nature* 380:240–243.
- See S (2000) Mutualism, parasitism and competition in the evolution of coviruses. *Philos Trans R Soc London Ser B* 355:1607–1613.
- Kerr B, Riley MA, Feldman MW, Bohannan BJ (2002) Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature* 418:171–174.
- Nowak MA, May RM (1992) Evolutionary games and spatial chaos. *Nature* 359:826–829.
- Durrett R, Levin SA (1994) The importance of being discrete (and spatial). *Theor Popul Biol* 46:363–394.
- Hassell MP, Comins HN, May RM (1994) Species coexistence and self-organizing spatial dynamics. *Nature* 370:290–292.
- Killingback T, Doebeli M (1996) Spatial evolutionary game theory: Hawks and Doves revisited. *Proc R Soc London Ser B* 263:1135–1144.
- Nakamaru M, Matsuda H, Iwasa Y (1997) The evolution of cooperation in a lattice-structured population. *J Theor Biol* 184:65–81.
- Eshel I, Sansone E, Shaked A (1999) The emergence of kinship behavior in structured populations of unrelated individuals. *Int J Game Theory* 28:447.
- Neuhauser C, Pacala S (1999) An explicitly spatial version of the Lotka-Volterra model with interspecific competition. *Ann Appl Prob* 9:1226–1259.

21. Szabo G, Hauert C (2002) Phase transitions and volunteering in spatial public goods games. *Phys Rev Lett* 89:118101.
22. Hauert C, Doebeli M (2004) Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature* 428:643–646.
23. Ohtsuki H, Hauert C, Lieberman E, Nowak MA (2006) A simple rule for the evolution of cooperation on graphs and social networks. *Nature* 441:502–505.
24. Santos FC, Pacheco JM, Lenaerts T (2006) Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proc Natl Acad Sci USA* 103:3490–3494.
25. Taylor PD, Day T, Wild G (2007) Evolution of cooperation in a finite homogeneous graph. *Nature* 447:469–472.
26. Hamilton WD (1964) The genetical behavior of social behavior I. *J Theor Biol* 7:1–16.
27. Dawkins R (1976) *The Selfish Gene* (Oxford Univ Press, Oxford).
28. Dawkins R (1982) *The Extended Phenotype* (Oxford Univ Press, Oxford).
29. Matteo JM, Johnston RE (2000) *Proc R Soc London* 267:695–700.
30. Sinervo B, et al. (2006) Self-recognition, color signals, and cycles of greenbeard mutualism and altruism. *Proc Natl Acad Sci USA* 103:7372–7377.
31. Lize A, et al. (2006) Kin discrimination and altruism in the larvae of a solitary insect. *Proc R Soc London Ser B* 273:2381–2386.
32. Riolo RL, Cohen MD, Axelrod R (2001) Evolution of cooperation without reciprocity. *Nature* 414:441–443.
33. Axelrod R, Hammond RA, Grafen A (2004) Altruism via kin-selection strategies that rely on arbitrary tags with which they coevolve. *Evolution (Lawrence, Kans)* 58:1833–1838.
34. Jansen VAA, van Baalen M (2006) Altruism through beard chromodynamics. *Nature* 440:663–666.
35. Hammond RA, Axelrod R (2006) Evolution of contingent altruism when cooperation is expensive. *Theor Pop Biol* 69:333–338.
36. Rousset F, Roze D (2007) Constraints on the origin and maintenance of genetic kin recognition. *Evolution (Lawrence, Kans)* 61:2320–2330.
37. Gardner A, West SA (2007) Social evolution: The decline and fall of genetic kin recognition. *Curr Biol* 17:R810–R812.
38. Hochberg ME, Sinervo B, Brown SP (2003) Socially mediated speciation. *Evolution (Lawrence, Kans)* 57:154–158.
39. Traulsen A, Nowak MA (2007) Chromodynamics of cooperation in finite populations. *PLoS ONE* 2:e270.
40. Levin SA, Segel LA (1982) Models of the influence of predation on aspect diversity in prey populations. *J Math Biol* 14:253–284.
41. Levin SA, Segel LA (1985) Pattern generation in space and aspect. *SIAM Rev* 27:45–67.
42. Axelrod R, Hamilton WD (1981) The evolution of cooperation. *Science* 211:1390–1396.
43. Nowak MA (2006) Five rules for the evolution of cooperation. *Science* 314:1560–1563.
44. Byrne D (1969) Attitudes and attraction. *Adv Exp Soc Psychol* 4:35–89.
45. Nahemow L, Lawton MP (1975) Similarity and propinquity in friendship formation. *J Personal Soc Psychol* 32:205–213.
46. Selfhout MHW, et al. (2007) The role of music preferences in early adolescents' friendship formation and stability. *J Adolesc* 32:95–107.
47. Tajfel H, Billig RP, Flament C (1971) Social categorization and intergroup behavior. *Eur J Soc Psychol* 1:149–178.
48. Burger JM, Messian N, Patel S, Prado AD, Anderson C (2004) What a coincidence! The effects of incidental similarity on compliance. *Personality Soc Psychol Bull* 30:35–43.
49. Rand DG, et al. (2009) Dynamic remodeling of in-group bias during the 2008 presidential election. *Proc Natl Acad Sci USA*, 10.1073/pnas.0811552106.
50. Moran PAP (1975) Wandering distributions and electrophoretic profile. *Theor Popul Biol* 8:318–330.
51. Kingman JFC (1976) Coherent random-walks arising in some genetic models. *Proc R Soc London Ser A* 351:19–31.
52. Wright S (1931) Evolution in Mendelian populations. *Genetics* 16:97–159.
53. Kingman JFC (1982) On the genealogy of large populations. *J Appl Prob* 19:27–43.
54. Wakeley J (2008) *Coalescent Theory: An Introduction* (Roberts & Company Publishers, Greenwood Village, CO).
55. Price GR (1970) Selection and covariance. *Nature* 227:520–521.
56. Rousset F (2003) A minimal derivation of convergence stability measures. *J Theor Biol* 221:665–668.