Electronic supplementary material

Appendix A: Repeated $n$-player Prisoner’s Dilemma for the smallest possible population size $N (N = n)$ for general strategy

Each game involves all the $N$ individuals in the population and consists of one or more rounds. We consider two strategies: ALLD and R. ALLD always defect no matter what other individuals do. R can be any strategy that cooperates at least once during each game when there are only one R and $N - 1$ ALLDs in the population. The payoffs, $V(C|l)$ and $V(D|l)$, have the same properties as above, given by (S1)-(S4).

First, let us consider whether strategy R is ESS$_N$ or not. From (S1) and (S3), we have $V(D|l) > V(C|l)$. This inequality means that a single mutant ALLD in a population of strategy R does not have a lower fitness and thus the first condition for R to be ESS$_N$ is violated. Therefore, we conclude strategy R that is ESS$_N$ does not exist.

Second, we examine whether ALLD is ESS$_N$ or not. Since $V(D|1) > V(C|1)$ and a mutant R cooperates at least once, selection opposes R invading ALLD, that is, the first condition for ALLD to be ESS$_N$ is satisfied. The fixation probability of R is given by

$$\rho_R = 1 \left( 1 + \sum_{k=1}^{N-1} \frac{1-w+wG_i}{1-w+WF_i} \right),$$

where $i$ denotes the number of R individuals and $F_i$ and $G_i$ are the expected payoffs of R and ALLD, respectively (see main text). Since $F_1 < G_1$ and $F_i \leq G_i$ ($2 \leq i \leq N - 1$) from the definition of R, the second condition for ALLD to be ESS$_N$ is satisfied. Therefore, we can conclude that ALLD is ESS$_N$.

To sum up, even in repeated $n$-player Prisoner’s Dilemma, cooperation cannot evolve when population size and group size are the same. Note that the argument in this
subsection applies for any payoff functions that satisfy (S1)-(S4): we do not have to assume specific payoff functions such \( V(C|k) = bk/n - c \) and \( V(D|k) = bk/n \). Also, the argument does not presuppose weak selection.

Appendix B: One-shot \( n \)-player Prisoner’s Dilemma for general population size

Groups of \( n \) individuals are sampled from the population of size \( N \) and a game is played within each group. A game consists of only one round, in each of which individuals either cooperate or defect. When some individuals cooperate, all the individuals in the group gain a benefit from it while only the cooperating individuals have to pay a cost.

Let \( V(C|l) \) and \( V(D|l) \) be the payoffs to individuals choosing cooperation and defection given that \( l \) of \( n \) individuals in the group choose cooperation. The \( n \)-player Prisoner’s Dilemma demands that these payoffs have the following properties:

\[
V(D|l) > V(C|l + 1), \quad (S1)
\]

\[
V(D|l + 1) > V(D|l), \quad (S2)
\]

\[
V(C|l + 1) > V(C|l), \quad (S3)
\]

\[
(l + 1)V(C|l + 1) + (n - l - 1)V(D|l + 1) > lV(C|l) + (n - l)V(D|l), \quad (S4)
\]

where \( 0 \leq l \leq n - 1 \). Using (S1) and (S3), we can conclude that D is always traditional ESS while C never is. Now, let us examine for general \( N \) whether either or both C and D can be ESS\(_N\). First, consider the case when a mutant D appears in a population of C. Using (S1) and (S3), we can obtain an inequality, that is,

\[
(N - 1)V(D|l) > (N - n)V(C|l + 1) + (n - 1)V(C|l).
\]

This inequality implies that the first condition for C to be ESS\(_N\) is violated (let \( a_k = V(D|k - 1) \), \( b_k = V(C|k) \) and \( l = n - 1 \) in (5) in the main text), that is, a single mutant D in a population of C has a higher fitness. Therefore, we conclude that C cannot be ESS\(_N\).
Second, suppose that a mutant C has appeared in a population of D. From (S1) and (S2), we obtain

\[(N - 1)V(C|l + 1) < (N - n)V(D|l) + (n - 1)V(D|l + 1),\]

which suggests that the first condition for D to be ESS\(_N\) is satisfied (let \(a_k = V(C|n - k + 1), b_k = V(D|n - k)\) and \(l = 0\) in (5) in the main text). The fixation probability of C is given by

\[
\rho_C = \frac{1}{1 + \sum_{k=1}^{N-1} \prod_{i=1}^{k-1} \frac{1 - w + wG_i}{1 - w + wF_i}},
\]

where \(i\) denotes the number of C individuals and \(F_i\) and \(G_i\) are the expected payoffs of C and D, respectively (see main text). From (S1) and (S2), we have

\[
\sum_{k=1}^{n} \binom{i-1}{n-k} \binom{N-i}{k-1} V(C|n - k + 1) < \sum_{k=1}^{n} \binom{i}{n-k} \binom{N-i-1}{k-1} V(D|n - k),
\]

that is, \(F_i < G_i\) \((1 \leq i \leq N - 1)\). Hence, \(\rho_C < 1/N\) and thus the second condition for D to be ESS\(_N\) is satisfied. Therefore, we can conclude that D is ESS\(_N\).

To sum up, in nonrepeated \(n\)-player Prisoner’s Dilemma, C is neither ESS\(_N\) nor traditional ESS, and D is not only ESS but also ESS\(_N\). Hence, cooperation cannot emerge in populations consisting of defectors. Note that the argument in this subsection applies for any payoff functions that satisfy (S1)-(S4) and for any selection intensity.

Appendix C: For \(n=2, 3\), no payoff matrix exists for which both A and B are traditional ESS but neither is ESS\(_N\).

In two-player games,

\[2a_1+a_2 < 2b_1+b_2\]

\[a_1+2a_2 > b_1+2b_2\]
\[ a_1 > b_1 \]
\[ a_2 < b_2 \]

All these inequalities cannot be satisfied simultaneously. Hence, for \( n=2 \), no payoff matrix exists for which both A and B are traditional ESS but neither is ESS\(_N\).

Similarly, in three player games,
\[ 3a_1 + 2a_2 + a_3 < 3b_1 + 2b_2 + b_3 \]
\[ a_1 + 2a_2 + 3a_3 > b_1 + 2b_2 + 3b_3 \]
\[ a_1 > b_1 \]
\[ a_3 < b_3 \]

All these inequalities cannot be met simultaneously. Hence, for \( n=3 \), no payoff matrix exists for which both A and B are traditional ESS but neither is ESS\(_N\).

Appendix D: Figure illustrates the relationship between payoff parameter \((b/c)\), group size \((n)\) and the minimum number of rounds \((m)\) required for \(\rho_{\text{ALLD}} < 1/N, \rho_{\text{TFT}_{n-1}} > \rho_{\text{ALLD}}\) or \(\rho_{\text{TFT}_{n-1}} > 1/N\) when \(N\) is large. The blue represents the number of rounds for which \(\rho_{\text{ALLD}} = 1/N\), while the green represents that for which \(\rho_{\text{TFT}_{n-1}} = \rho_{\text{ALLD}}\) and the red represents that for which \(\rho_{\text{TFT}_{n-1}} = 1/N\). The parameter space is divided into the following four regions: (i) TFT\(_{n-1}\) is ESS\(_N\) and ALLD is not ESS\(_N\); (ii) both TFT\(_{n-1}\) and ALLD are ESS\(_N\) and \(\rho_{\text{TFT}_{n-1}} > \rho_{\text{ALLD}}\); (iii) both TFT\(_{n-1}\) and ALLD are ESS\(_N\) and \(\rho_{\text{TFT}_{n-1}} < \rho_{\text{ALLD}}\); and (iv) TFT\(_{n-1}\) is not ESS\(_N\) and ALLD is ESS\(_N\). This figure reduces to figure 1a when \(b/c = 1.5\) and to figure 1b when \(b/(nc) = 0.51\).