

Supporting Information

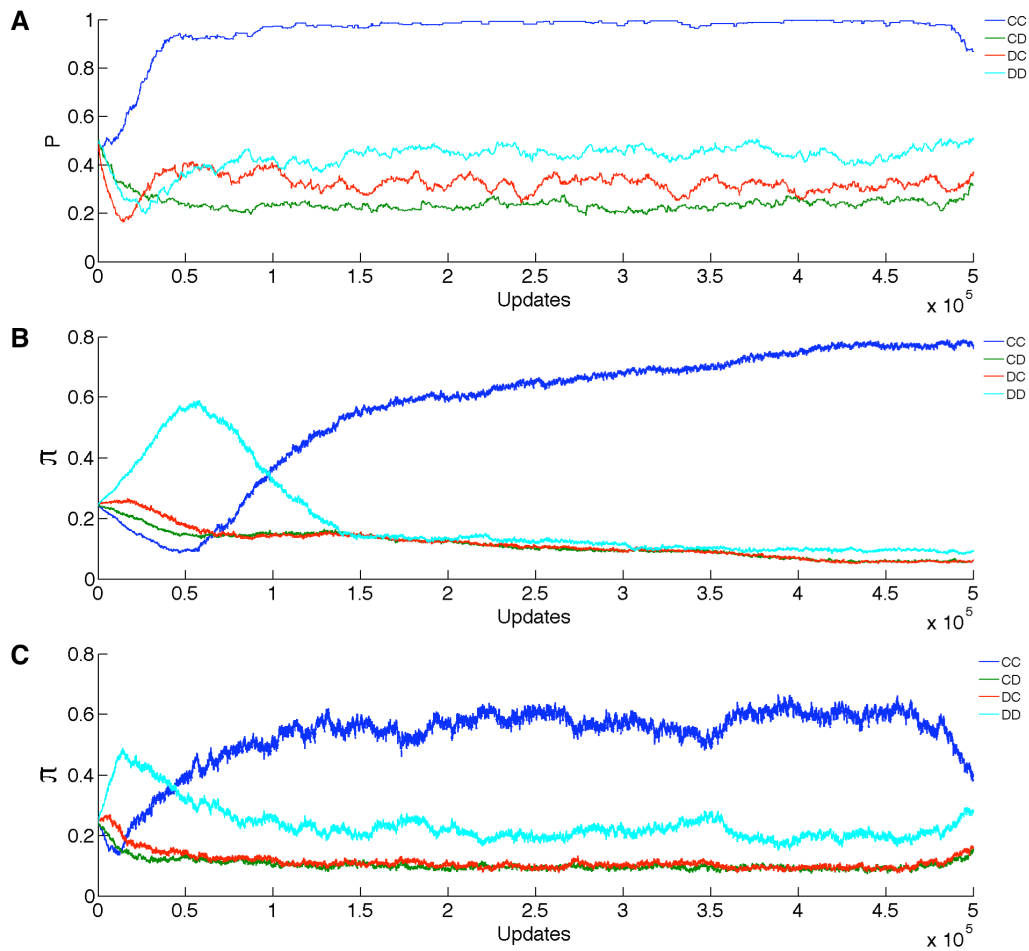


Figure S1. Average gene probabilities and play statistics. LOD gene probabilities (P_{XY}) and play statistics (π_{XY}) for a spatially-structured population, averaged over 80 experiments (500,000 updates each), at different μ and fixed r (1%). P_C and π_C are omitted because P_C drifts almost neutrally (see Methods) **A**, Average gene probabilities recorded at $\mu=1\%$. **B**, Play statistics recorded at $\mu=0.1\%$. **C**, Play statistics recorded at $\mu=2\%$.

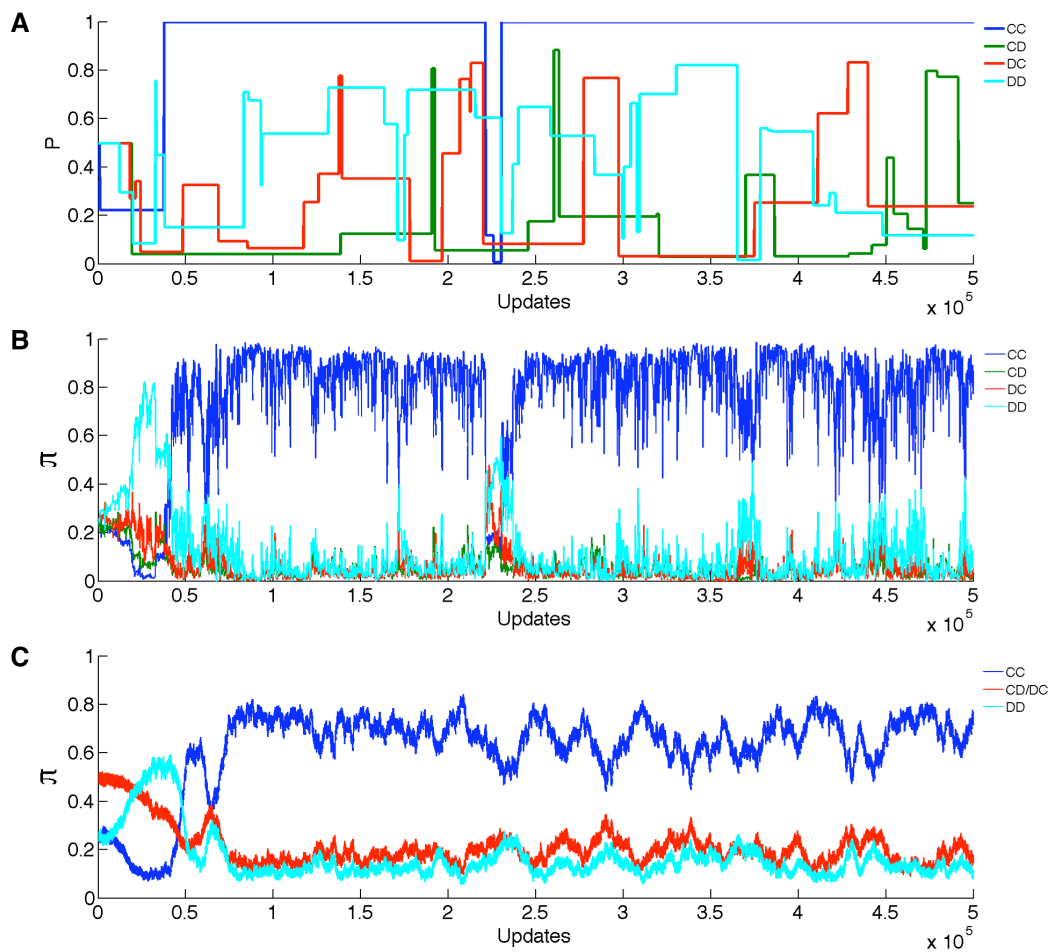


Figure S2. LOD and population genotypes and phenotypes. Single run LOD gene probabilities (P_{XY}) and play statistics (π_{XY}), as well as population average play statistics, for a spatially-structured population at $\mu=1\%$ and $r=1\%$. P_C and π_C are omitted because P_C drifts neutrally (see Methods). **(A)** LOD gene probabilities. **(B)** LOD play statistics. **(C)** Average population play statistics.

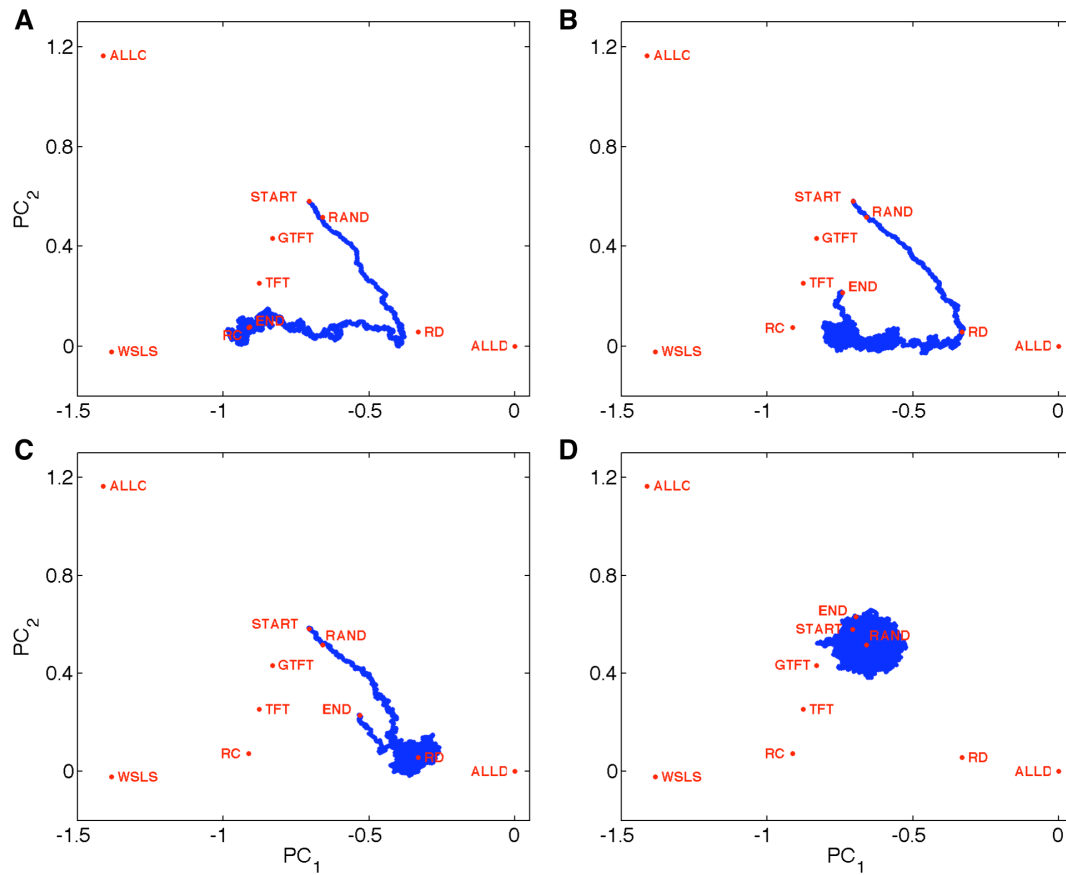


Figure S3. Evolutionary trajectories and attractors for well-mixed populations. All trajectories start at the same point ('START'), and move towards the strategy marked by 'END'. Several well-known strategies provide landmarks in strategy space: 'TFT': $(P_{CC}, P_{CD}, P_{DC}, P_{DD})=(1,0,1,0)$, 'ALLC'= $(1,1,1,1)$, 'ALLD'= $(0,0,0,0)$, WSLs= $(1,0,0,1)$, GTFT= $(1,0.333,1,0)$, START= $(0.5,0.5,0.5)$. All experiments shown are run at replacement rate $r=1\%$ for well-mixed populations. **(A)**, Evolution of the average LOD for $\mu=0.5\%$. RC marks the consensus genotype of this trajectory (described in supplementary text). This attractor is not the same as 'END' because that genotype lies past the most recent common ancestor of the population. **(B)** Trajectory for $\mu=1.5\%$, close to the critical mutation rate. **(C)** Trajectory for $\mu=5\%$. 'RD' marks the consensus genotype for these parameters. **(D)** Trajectory for $\mu=50\%$. 'RAND' marks the consensus genotype for these parameters.

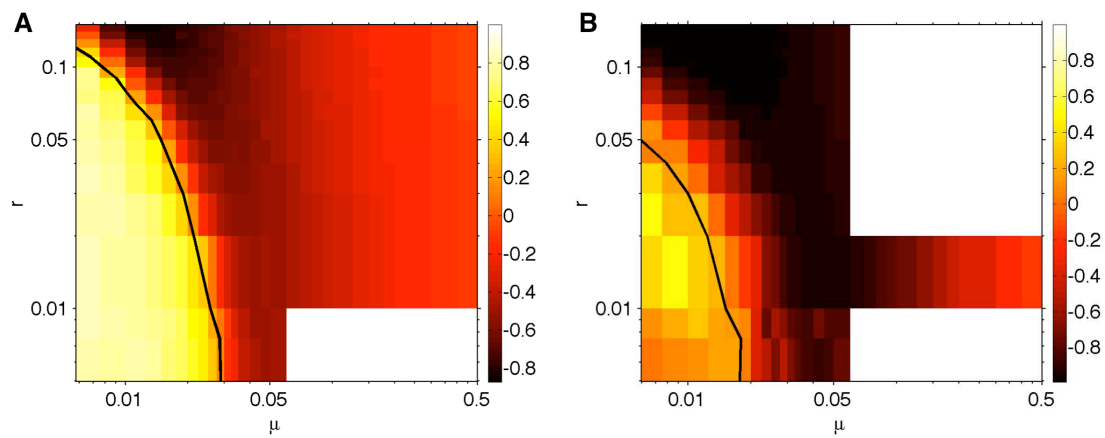


Figure S4. Quantitative phase transition diagrams as a function of μ and r . Coloring is applied according to the order parameter (m) with dark red to black indicating defection ($m < -0.2$), light yellow to light orange indicating cooperation ($m > 0.2$) and orange indicating a transition regime of equal cooperation and defection ($-0.2 \leq m \leq 0.2$). White colored areas contain no recorded data. **(A)** Spatially-structured. **(B)** Well-mixed.

Text S1. Description of consensus strategies.

At low mutation rates (<1%) and a fixed replacement rate of 1%, robust cooperation quickly emerges as the dominant strategy for both well-mixed and spatially-structured populations. Robustness is generally regarded as a measure of a system's performance in the face of perturbation or uncertainty. In this context, robust cooperation describes three basic behaviors that allow players to play well against defective players and even when faced with their own strategy, described in the main text.

The consensus genotypes at high mutation rates (5%) and a fixed replacement rate of 1%, for both population types, are representative of robust defection. Robustness in this case is less of a mandate in a cooperative regime, since cooperation is inherently a

more risky behavior in PD. Robust defectors have to maintain D exchanges while at the same time be willing to bait cooperative strategists with C plays and exploit them in the process. A P_{CC} probability close to 0.5 with a bias towards defection compared to the P_{CC} of RC is indicative of the unwillingness to engage in extended cooperative exchanges but at the same time a willingness to establish limited CC exchanges in hopes of future exploitation of a cooperative strategist. On the other hand, a very low P_{DD} probability expresses eagerness to maintain defective play. For the well-mixed population, the P_{DD} probability is much smaller than the equivalent probability in the spatially-structured population, which might be due to the absence of clusters under well-mixing giving rise to a higher degree of uncertainty experienced by players. The same reasoning might be applicable to the low P_{DC} probability. Defector strategists in well-mixed populations are more willing to take advantage of cooperative players since they play against many more such players due to the absence of cooperative cluster shielding.

Text S2. Experimental statistics

For the spatially-structured phase transition experiments a total of 532 pairs of different mutation and replacement rate experiments were run, for a minimum of 13 and a maximum of 80 replicates for each pair (depending on the replacement rate), leading to 14,576 replicates in total. For the well-mixed phase transition experiments a total of 304 pairs of different mutation and replacement rate experiments were run, for a minimum of 10 and a maximum of 80 replicates each, leading to 10,080 replicates in total. For both population types, each experiment was run for 500,000 updates leading to approximately 7.3 billion updates for spatially-structured and 5 billion updates for well-mixed populations. All experiments were run in parallel on up to 25 dual core 2.6GHz (Intel® Pentium® Processor E5300) computers for an approximate total of about 8,500 hours for the spatially-structured and 5,900 hours for well-mixed population

experiments. Data collected from the phase transition experiments were used to generate Figures 1, 2 and 3 as well as figures S1, S2, S3 and S4. For Figure 4A and memory depth 2, 16 different mutation rates were run at 1% replacement rate (80 replicates each), for a total of 1,280 experiments requiring 1,280 CPU hours. For Figure 4B, bits 1 to 5 were run for 40 replicate experiments at 15 different mutation rates (and a fixed replacement rate of 1%), for a total of 3,000 experiments (1,000 CPU hours). In total, the study required almost 29,000 individual experiments requiring a total of 1.9 CPU years (about 10 CPU weeks in parallel).

Table S1. Consensus genotypes for different mutation rates and population structures. Mean probabilities for each gene averaged over 80 average LODs, with variance in brackets. SS: spatially-structured population, WM: well-mixed population, COOP: cooperator, DEFEC: defector.

	P_C	P_{CC}	P_{CD}	P_{DC}	P_{DD}
SS COOP ($\mu=1\%$)	0.647 (0.088)	0.989 (0.005)	0.234 (0.035)	0.318 (0.075)	0.448 (0.054)
SS DEFEC ($\mu=5\%$)	0.481 (0.084)	0.458 (0.091)	0.315 (0.062)	0.243 (0.073)	0.325 (0.064)
WM COOP ($\mu=1\%$)	0.595 (0.098)	0.893 (0.056)	0.247 (0.038)	0.247 (0.079)	0.356 (0.076)
WM DEFEC ($\mu=5\%$)	0.442 (0.081)	0.460 (0.084)	0.325 (0.059)	0.063 (0.018)	0.053 (0.012)