

# Community Annotation and the Evolution of Cooperation: How Patience Matters

Atin Basuchoudhary<sup>1,\*</sup>, Vahan Simoyan<sup>2</sup> and Raja Mazumder<sup>3</sup>

<sup>1</sup>Department of Economics, Virginia Military Institute, Lexington, VA 24450, USA

<sup>2</sup>Food and Drug Administration, CBER, 1451 Rockville pike, Rockville, MD, 20850, USA

<sup>3</sup>Department of Biochemistry and Molecular Biology, George Washington University, Washington, DC 20037, USA

**Abstract:** We investigate why biologists fail to contribute to biological databases although almost all of them use these databases for research. We find, using evolutionary game theory and computer simulations, that (a) the initial distribution of contributors who are patient determines whether a culture of contribution will prevail or not (b) institutions (where institution means “a significant practice, relationship, or organization in a society or culture”) that incentivize patience and therefore limit free riding make contribution more likely and, (c) a stable institution, whether it incentivizes patience or not, will increase contribution. As a result we suggest there is a trade-off between the benefits of changing institutions to incentivize patience and the costs of the change itself. Moreover, even if it is possible to create institutions that incentivize patience among scientists such institutions may nevertheless fail. We create a computer simulation of a population of biologists based on our theory. These simulations suggest that institutions should focus more on rewards rather than penalties to incentivize a culture of contribution. Our approach therefore provides a methodology for developing a practical blueprint for organizing scientists to encourage cooperation and maximizing scientific output.

**Keywords:** Community annotation, Cooperation, Biocuration.

## INTRODUCTION

Science is a cooperative process. Cooperation, is an action which incurs a cost  $c$  to the individual that performs it, and provides a benefit  $b$  to the recipient of that action [1]. Researchers build off the work of others. Thus, if researchers do not have access to the work of others, scientific progress slows. Such access is possible if, for example, biologists were to contribute annotations to public databases. This process of contributing annotations is also known as community annotation. Researchers can then access the databases and use annotations contributed by others. Thus, the level of contributions capture how much biologists are willing to cooperate with each other without any expectation of co-authorship. Of course each biologist reaps *private* benefits from using these databases. The question then is do biologists actually cooperate by contributing to these databases? Unfortunately, they do not [2].

There have been several attempts to engage bench scientists in contributing annotations to scientific databases. Several articles describe the huge potential of such an approach [3-15]. In fact, the scientific community hardly ever contributes by providing annotations (such as protein/gene function) to scientific databases even though they use the biological data available in these databases quite

intensively [2]. Fewer than 1% of all possible contributors actually contribute [16]. Scientists have hypothesized that community annotation has not been successful because of technical barriers [17]. However, several databases provide simple mechanisms for contributing annotations. Thus, it appears that the problem of non-contribution is social rather than technological.

In the social sciences it is well known that patient people (people who are willing to contribute or act today and willing to wait to reap benefits later) cooperate while impatient people do not [18]. This idea is so well known that it is unclear who came up with it – hence its name, the folk theorem [19]. The folk theorem states that any feasible payoff profile that strictly dominates the minmax profile can be realized as a Nash equilibrium payoff profile, with sufficiently large discount factor. People with a high discount factor value future payoffs more. That is, they are willing to wait for the future. Thus, according to the folk theorem, patient biologists, i.e. biologists with a high discount factor, should cooperate by contributing to databases [20]. Why don't biologists cooperate by contributing to these databases? The Folk theorem states that biologists (or anybody for that matter) do not cooperate because they are impatient. So more cooperation may be possible if biologists are patient. How might patience evolve? We suggest a theory.

The Merriam Webster dictionary defines culture as “the integrated pattern of human knowledge, belief, and behavior that depends upon the capacity for learning and transmitting knowledge to succeeding generations.” Research in

\*Address correspondence to this author at the Department of Economics, Virginia Military Institute, Lexington, VA 24450, USA; Tel: 202-994-5004; Fax: 202-994-8974; E-mail: [BasuChoudharyA@vmi.edu](mailto:BasuChoudharyA@vmi.edu)

behavioral economics has established the psychological roots of patience [21]. Thus, patience – or the lack thereof – is ingrained in human behavior. Moreover, people differ in how patient they are [22]. Since human behavior can be culturally influenced and patience is part of human behavior, then it seems likely that patience can be culturally influenced as well [23]. Thus people can learn to be more or less patient from others when patience levels vary. In other words, there is diversity in the patient trait and a *cultural* transmission process for it. The mathematics of natural selection ensures that fitter cultural traits survive in the population while less fit traits perish. This process leads to a stable selection of traits. Thus an evolutionary game theory approach may help us understand whether a culture of patience can be evolutionarily stable in a population of biologists. In other words, we apply the mathematical tools that explain biological evolution to understand a process of cultural evolution.

We use a theoretical model [24] and extend it. In this model we assume that people in the patient culture cooperate with others while people in the impatient culture do not. Then, we hypothesize that a culture of patience is “infectious” in a way that is amenable to analysis with evolutionary game theory. This is a major departure from the theoretical and empirical literature on patience [25-27]. Indeed, in spite of evidence to the contrary [22], experiments that deal with the evolution of cooperation largely ignore the variability of patience in a population [25, 26]. Our innovation lies in proposing a model for the evolution of patience as a cultural trait. Recall that the Folk Theorem suggests that cooperation is possible if people are patient enough. Thus by establishing a process for the evolution of patience we can also address the evolution of cooperation. This approach also provides a new kind of framework for designing experiments with human subjects to test how cooperation may evolve in a population.

Our fundamental argument is as follows. Patient people tend to cooperate [18, 19]. People are heterogeneous in patience [21, 22]. We assume that this heterogeneity is cultural [23]. These cultures are subject to selection pressures from the social environment and may therefore evolve [24]. The conditions that allow patience to evolve toward a stable equilibrium will also increase cooperation because of the folk theorem. In this paper we derive a theory for how different parameters influence the evolution of patience and therefore cooperation. We then run computer simulations of a population of agents, each programmed with different frequencies of contribution. The agents are distributed according to different probability distribution functions. These programmed agents are then allowed to interact with each other randomly. A *generic* replicator dynamic is applied to these interactions. Since agents in the simulation are not hardwired to follow the path directed in the model it is not clear *a priori* whether the population will behave in a way predicted by the model or not. We find that they do.

Our evolutionary approach suggests that even when the benefits of patience (cooperation) are obvious there are circumstances where impatience (and therefore non-cooperation) can become a stable cultural norm – and therefore explain why biologists do not cooperate with each other by contributing annotations to databases. This study

suggests ways to *change* the circumstances that make cooperation less likely. Our computer simulation tests whether cooperation can indeed become a cultural norm as these circumstances change.

Contributing to molecular biology databases has no immediate reward for the contributor, they have to wait for others to contribute to take advantage of new information. In fact, contributors may benefit only if others contribute as well. Thus, there are clear benefits from cooperating by contributing. This is certainly true for systems biology research where information on a large number of proteins from a proteome is necessary to develop any reasonable testable hypothesis. Currently, the majority of experimental data available in publications is not available in online databases. For example, only 10% of available publications on *Bacillus subtilis* (a well studied model organism) are cited in UniProtKB/Swiss-Prot ([www.uniprot.org](http://www.uniprot.org)). Using the search term *Bacillus subtilis* [MeSH Terms] or “*Bacillus subtilis*” at [pubmed.org](http://pubmed.org) retrieves 20,000 – 25,000 publications. A UniProtKB/Swiss-Prot search of publications (including mapped pubmed IDs from non-UniProt databases such as Model Organism Databases) associated with *Bacillus subtilis* (including all sub-species) retrieves 4,286 Swiss-Prot entries with only 2,310 unique publications (UniProtKB release 2010\_05). If one also considers supplementary tables, it is clear that a large part of the knowledge available on *Bacillus subtilis* is not accessible in a structured format that would allow integrated analysis. This lack of information in molecular biology databases when compared to what is available in publications is true for all species and other databases. This happens because it takes an immense amount of time and effort to annotate a specific entry by a biocurator [28, 29]. It is true that the workflow of biocurators at UniProtKB does not include reading all papers or even citing all the papers that they read unless the paper has relevant information that can be curated, but it is clear that outside of the paid biocurators’ efforts there is very little information in UniProtKB annotations that has been submitted by the community, even though the community uses it heavily and there are easy to use submission forms on the website. Further, as mentioned above, the biological community may not be contributing because the benefits of such contributions in terms of hypothesis generation and research are uncertain. Moreover, the benefits may accrue to researchers other than the contributor. In addition, these benefits, given the lag time for research, may accrue well in the future. The social interaction among biologists thus has two connected elements. First, biologists can use the contributions of others in their own research without contributing to the databases – free ride off the work of others in the parlance of the social sciences. Second, biologists have an incentive to get an immediate gain – getting publications in time for tenure decisions or writing a grant application to pay current salary for example – rather than wait to get benefits from longer term projects. In other words, impatient biologists would not contribute to databases while patient biologists will.

We model a society with two cultures, patient and impatient. Biologists “inherit” or choose to join one or the other culture. Biologists in either culture interact in repeated random pairings [30]. We then find the evolutionary stable strategies (ESS), i.e. the strategy or culture that, if adopted

Table 1. The Evolutionary Stage Game

	Patient ( $r_L$ )	Impatient ( $r_H$ )
Patient ( $r_L$ )	$\frac{e}{2} \left( \frac{1+r_L}{1+r_L-p} \right), \frac{e}{2} \left( \frac{1+r_L}{1+r_L-p} \right)$	$0, \alpha e \left( \frac{1+r_H}{1+r_H-p} \right)$
Impatient ( $r_H$ )	$\alpha e \left( \frac{1+r_H}{1+r_H-p} \right), 0$	$\frac{e\alpha}{2} \left( \frac{1+r_H}{1+r_H-p} \right), \frac{e\alpha}{2} \left( \frac{1+r_H}{1+r_H-p} \right)$

by all members of a society, cannot be “invaded” by any alternative cultures [31]. Thus, we find the conditions under which impatience or patience can be stable cultures. We recognize that there may be many patterns of interactions – and simulate some of those in our computer model. However, our modeling approach, where random pairings lead to the transmission of cultural phenotypes, is simplest. We also recognize that patience heterogeneity need not be binary. However, assuming binary traits makes the mathematics much simpler without losing any of the fundamental features that we wish to focus on.

Institutions (where institution means “a significant practice, relationship, or organization in a society or culture”) that incentivize patience should be critical in making contribution to the dominant culture in biology. Further, the perceived longevity of an institutional mechanism should matter as well. Our theoretical model characterizes these sorts of dynamics. We find that our results are robust to a less restrictive distribution of contributory behaviors when we simulate this society of biologists in a computer generated “realistic” environment. Thus, we suggest that our methodological approach could provide the policy framework for funding agencies and academic institutions to organize scientific effort optimally.

## METHODS

Evolutionary games study the evolution of strategies or behaviors in a population. Players receive payoffs that are interpreted as the “fitness” of their respective *cultures* (patience vs. impatience). Thus effectively cultures or strategies interact with each other and not players. In other words, in evolutionary game theory, strategies, and not players, are subject to selection pressures. The fitter culture’s proportion increases in a population.

We use the general version of the models introduced by Basuchoudhary, Allen and Siemers [24]. We model biologists as belonging to one of two possible cultures. In one culture biologists contribute to a database because they are willing to wait for greater future benefits. In the other culture biologists do not contribute because they seek immediate rewards. Biologists who contribute to the database are **patient** relative to biologists who do not contribute. Thus, patience is the source of cultural heterogeneity in our model. This binary cultural trait is represented by the discount rate  $r$ . The discount rate is an interest rate used to calculate the present value of future gains. Impatient people discount future gains more heavily than patient people. Thus, biologists in a patient culture have a lower discount rate ( $r_L$ ) than biologists in the **impatient** high discount rate ( $r_H$ ) culture. This sort of social interaction has the structure of a coordination game. Contributors in our

stylized model only benefit if others contribute as well while non-contributors can free ride off the contributions of others.

In our model fitness depends on certain parameters. The total fitness when biologists from the contributing or patient culture interact,  $e$ , defines the efficient outcome for the society of biologists as a whole. Therefore  $e$  represents a “good” outcome with long term scientific innovation. We assume that biologists receive the full benefit of their cooperation in contributing to the database. This assumption allows us to abstract away from principal-agent problems that may arise. In addition,  $\alpha \in [0, \frac{1}{2}]$  tracks the proportion of the efficient outcome  $e$  appropriated by biologists in the non-contributing culture. Thus  $\alpha$  tracks the extent to which the non-contributing or impatient culture encourages “free riding” off the contributions of others. Social institutions reflect culture. Thus,  $\alpha$  is a way of capturing the effect of social institutions on the incentive to free ride. Please note that we explicitly limit the range for  $\alpha$  so as to limit the strategic interaction investigated in this paper to an assurance game. For example, for any  $\alpha \in (\frac{1}{2}, 1)$  the game structure changes to a prisoners dilemma. We, however, believe that a contributory culture of biologists’ is best specified by  $\alpha \in [0, 1/2)$  since there are increasing returns to contributions when everyone coordinates on contributing to a database [32].

In addition to these parameters we add  $p$  – the probability the game continues. Again, recall that the standard definition of institutions includes the rules of the game. Thus  $p$  tracks the level of uncertainty in the institutional environment of research. Thus, a falling  $p$  might indicate an increased likelihood of an end to the current scientific funding regime; i.e. an increased likelihood of institutional change. For example, say National Science Foundation (NSF) funding is driven by a purely democratic process with no input from scientists. This process may change the kinds of projects that are funded. We assume that  $p$  is independent of history.

The parameters described above,  $\alpha, r_H, r_L, e$ , and  $p$  give us the stage game described in Table 1. The table represents the fitness of each strategy relative to each other [33, 34]. We assume that pairs of biologists from the patient and impatient cultures interact with each other for long periods of time. A *cultural* replicator dynamic process transmits behavior from biologist to biologist. We do not explicitly model this transmission process. However, mimicry is sufficient for evolutionary models to work [35]. This modeling approach simplifies the mathematics and is quite standard in evolutionary game theory [36]. If biologists from the patient culture interact the fitness of the patient culture is

$e/2$  in each period. However, over time the present value of this interaction is  $\frac{e}{2} \left( \frac{1+r_L}{1+r_L-p} \right)$ . Notice that appropriate discount factor is  $r_L$  since we are calculating the fitness of the patient culture. If one biologist is from a patient culture and another biologist is from an impatient culture then the patient biologist receives a fitness of 0 while the impatient biologist gets a fitness of  $\alpha e$ . In other words, the patient biologist does not reap the rewards of her contribution while the impatient biologist benefits from the patient biologists' contribution. The present value of this interaction is 0 for the patient biologist and  $\alpha e \left( \frac{1+r_H}{1+r_H-p} \right)$  for the impatient biologist. Notice that the appropriate discount factor is  $r_H$  since we are calculating the fitness of the impatient culture. If both biologists are from the impatient culture then neither contribute. Therefore, in this case both impatient biologists receive a fraction  $\alpha/2$  of the efficient output of science  $e$ . Thus the present value of the fitness of the impatient culture when it interacts with another impatient culture is  $\frac{e\alpha}{2} \left( \frac{1+r_H}{1+r_H-p} \right)$ .

Note that for any  $\alpha$  that takes a value between 0 and 1/2 the highest possible fitness arises when people from a patient (cooperative and therefore contributing) culture interact with other people from the patient culture. However, only a proportion  $x$  of the population has a culture of patience. Thus, there is a chance  $(1-x)$  that patient biologists may interact with impatient (non-cooperative and therefore non-contributing) biologists. Note further that the parameter restrictions on  $\alpha$  suggest that the overall fitness of a world where impatience (and therefore non-cooperation and non-contribution) prevails is less than the overall fitness of a world where patience (and therefore cooperation and contribution) is the norm. This, perhaps, should not be surprising to the reader. However, overall, if a particular culture has an expected payoff above the average fitness of the entire population then the percentage of biologists belonging to that culture will increase. And of course, overall, if a particular culture has an expected fitness below the average fitness of the entire population then the percentage of biologists belonging to that culture will decrease. In other words, the mathematics of replicator dynamics drive the evolution of culture in our model.

We use a replicator dynamic approach to solve for the ESS. An ESS is stable in the sense that small changes, i.e. small proportions of a population playing the non-ESS strategy, cannot invade a population successfully. This is a static concept. A replicator dynamic approach suggests that an ESS is stable precisely because it arises out of the dynamics of a process that generates the fittest strategy [31]. We find that the success of a culture of patience depends on whether the actual proportion of the biologist population belonging to the patient culture exceeds some critical level  $x^*$ .

This approach allows us to derive three propositions. The propositions and the derivations follow. We discuss these propositions in the Discussion section.

Proposition 1. [There are no stable equilibria where both cultures coexist.  $x^* = \frac{1}{\alpha \left( \frac{1-\delta_H}{1-\delta_L} \right) - 1}$  is a tipping point. Thus,

patience, the extent to which institutional structures incentivize free riding, the probability that these institutional structures persist, and the initial proportion of contributors in a population of biologists determine whether a culture of contribution will prevail or not].

**PROOF OF PROPOSITION 1**

The expected payoff from the Patient ( $r_L$ ) strategy is

$$E(r_L) = \frac{xe}{2} \left( \frac{1+r_L}{1+r_L-p} \right) \tag{1}$$

and from Impatient ( $r_H$ ) strategy is

$$E(r_H) = \alpha ex \left( \frac{1+r_H}{1+r_H-p} \right) + \frac{\alpha e}{2} (1-x) \left( \frac{1+r_H}{1+r_H-p} \right) \tag{2}$$

The patient and impatient strategy provide the same expected payoff when  $E(r_L) = E(r_H)$ , i.e.,

$$\frac{xe}{2} \left( \frac{1+r_L}{1+r_L-p} \right) = \alpha ex \left( \frac{1+r_H}{1+r_H-p} \right) + \frac{\alpha e}{2} (1-x) \left( \frac{1+r_H}{1+r_H-p} \right)$$

or

$$\frac{xe}{2} \left( \frac{1+r_L}{1+r_L-p} \right) = \left( \frac{1+r_H}{1+r_H-p} \right) \left( \alpha ex + \frac{\alpha e}{2} (1-x) \right)$$

or

$$\frac{xe}{2} \left( \frac{1+r_L}{1+r_H} \right) \left( \frac{1+r_H-p}{1+r_L-p} \right) = \left( \alpha ex + \frac{\alpha e}{2} (1-x) \right) \tag{3}$$

Define  $\delta_L = p/(1+r_L)$  and  $\delta_H = p/(1+r_H)$ . Thus,

$$\frac{1+r_L}{1+r_H} = \frac{\delta_H}{\delta_L} \tag{4}$$

and,

$$\frac{1+r_H-p}{1+r_L-p} = \frac{\frac{1}{\delta_H}-1}{\frac{1}{\delta_L}-1} \tag{5}$$

Substituting (4) and (5) into (3) gives us

$$\frac{xe}{2} \left( \frac{1-\delta_H}{1-\delta_L} \right) = \left( \alpha ex + \frac{\alpha e}{2} (1-x) \right)$$

or,

$$x^* = \frac{1}{\frac{1}{\alpha} \left( \frac{1-\delta_H}{1-\delta_L} \right) - 1} \tag{6}$$

Thus for the patient strategy to be preferred over the impatient strategy

$$x > x^* = \frac{1}{\frac{1}{\alpha} \left( \frac{1-\delta_H}{1-\delta_L} \right) - 1} \tag{7}$$

Proposition 2. [The likelihood that the culture of non-contribution will spread through the population of biologists increases with the extent to which institutional structures incentivize free riding].

**PROOF OF PROPOSITION 2**

Notice in equation (6) that as  $\alpha$  rises the denominator of  $x^*$  becomes smaller. Thus, as  $\alpha$  rises so does  $x^*$ .

Proposition 3. [An increase in the likelihood that the current institutional structure will continue makes it easier for a culture of contribution to take root since  $\frac{dx^*}{d\alpha} < 0$ ].

**PROOF OF PROPOSITION 3**

Proposition 3 states that  $x^*$  is decreasing in  $p$ . To demonstrate this, we must show that the derivatives of  $x^*$  with respect to  $r_H$  and  $p$  are always negative. Recall from (7) that after substituting  $\delta_H$  and  $\delta_L$ ,

$$x^* = \frac{1}{\alpha \left( \frac{1 - \frac{p}{1+r_H}}{1 - \frac{p}{1+r_L}} \right) - 1}$$

We take the term  $\left( \frac{1 - \frac{p}{1+r_H}}{1 - \frac{p}{1+r_L}} \right)$  from the denominator and take the first derivative with respect to  $p$  to find that  $\frac{d}{dp} \left( \frac{1 - \frac{p}{1+r_H}}{1 - \frac{p}{1+r_L}} \right) = \left( \frac{r_H - r_L}{r_H + r_L} \right) \left( \frac{r_L + 1}{(r_L - p + 1)^2} \right)$ . Notice,  $(r_H - r_L) > 0, r_H \geq 0, r_L \geq 0$

by definition. Also,  $\left( \frac{r_L + 1}{(r_L - p + 1)^2} \right) > 0$ . Thus,  $\left( \frac{r_H - r_L}{r_H + r_L} \right) \left( \frac{r_L + 1}{(r_L - p + 1)^2} \right) > 0$ . In other words, the denominator of  $x^*$  rises with  $p$ . Thus  $x^*$  itself falls as  $p$  rises.

This concludes the proofs for our propositions.

Propositions 1, 2, and 3 above, show how changes in the parameters that affect the fitness of cultures in a population of biologists affect the likelihood of a culture of contribution taking root. In fact,  $\alpha, p$ , and  $r_H$  – the variables of interest in those three propositions – affect the rewards and penalties from pursuing a culture of contribution (the patient culture) relative to the rewards and penalties from pursuing a culture of non-contribution (the impatient culture). However, this process may take a long time. Thus, finding evidence for this process requires us to follow large groups of people for long periods of time or design expensive experiments with human subjects. Therefore, as an initial attempt to check whether our results are robust, we use a computer model to simulate a society of biologists in order to both test the robustness of our mathematical results as well as get a sense of the sensitivity of our predictions to changes in the rewards and penalties of contributing. In this simulation, we specify a replicator dynamic process based on various distributions of several species or cultures of biologists. The *frequency of contribution* is the source of heterogeneity among cultures in our simulation. Thus, our computer simulations do not “hardwire” patience in any way. We prefer instead to see if the predictions of our theoretical model are borne out when the level of contribution is the only choice made by “biologists” in our simulation. We do this to avoid the possibility of a simulation that gets our results because our results are already pre-programmed in the simulation. However, as biologists with different levels of contribution

interact, the order of the fitness of these interactions follows the payoff structure in Table 1. We therefore suggest that if biologists in the simulation behave predictably like biologists in our theoretical model then the simulation provides evidence in favor of our theory.

Our computer models simulate population dynamics to test our results. To perform these population dynamics simulations we have adopted a basic Metropolis technique to devise an efficient Monte-Carlo scheme for a sampling controlled by a generic continuous replicator dynamics equation  $\dot{x} = x_i(f_i(x) - \sum x_i f_i(x))$ . Here  $i$  is the index of the sub species of the population and  $f_i$  denotes the fitness of the  $i^{th}$  sub species while  $\sum x_i f_i(x)$  is the average fitness of the population. This method, therefore, first constructs a population with  $m$  members by randomly assigning frequency of contributions  $c_m$  from a given distribution. In reality, and unlike our theoretical model, people will not be either contributors or non-contributors. They are more likely to contribute along a spectrum of frequencies. Thus, some members of our simulated population contribute more frequently. These members of the simulated population are distributed according to a probability distribution. We report the results from two initial distributions, a uniform distribution, and a normal distribution. These distributions represent assumptions about how often biologists contribute (see Fig. 1). The broad results reported below are robust to the choice of distribution.

The method then attempts Monte-Carlo procedure of creation, deletion, and mutation (displacement) during large number of generations. One of two methods subsequently determines the randomized probability of acceptance for a certain operation: species derivative by replicator dynamics

$$rand() == x_i(f_i(x) - \sum x_i f_i(x)) \tag{8}$$

and normalized cumulative distribution function of the mean centered payoff,

$$rand() == \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(f_i(c_i) - \frac{\sum f_i(c_i)}{N})^2}{2\sigma^2}} \tag{9}$$

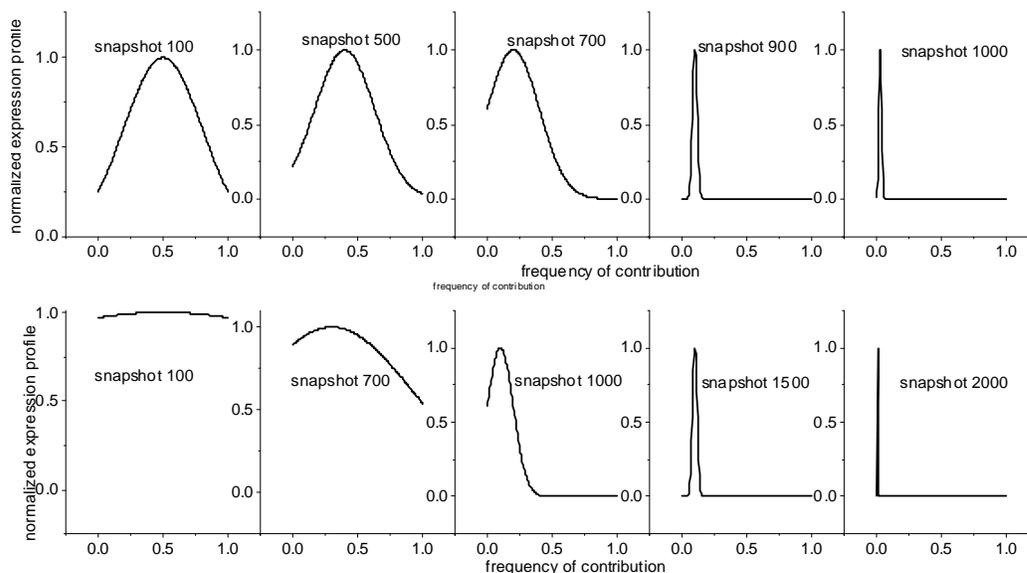
It is important to note that both equations yield the same equilibrium for binary systems, but the second equation allows for ESS distributions with more than two multiple quasi-species concurrently existing.

Analytical evaluation of continuous mode replicator dynamics leads to the following equation

$$\dot{x} = c(x) * [x * CC * M + x * CD * (1 - M) + (1 - x) * DC * M + (1 - x) * (1 - M) * DD] \tag{10}$$

where  $CC, CD, DC, DD$  are the coefficients in payoff matrix (for the simulations we started from  $CC=3, CD/DC=0/2, DD=1$ ).  $M$  is the total momentum of frequency distribution function

$$M = \int_0^1 x * c(x) dx \tag{11}$$



**Fig. (1).** Series of expression profile snapshots taken during population dynamics simulations for normal (top diagrams) and uniform (bottom diagrams) distributions. Abscissa (x-axis) shows the frequency of contribution by a subspecies and ordinate (y-axis) show the scaled probability of finding a member with such characteristics. Both panels show that starting from non-biased normal (extreme top left panel figure) or uniform (extreme bottom left panel figure) distributions result in completely non-contributing populations. Top panel shows that a population that has a preferred culture of non-contribution converges quicker to a total non-contributing population while the bottom panel (without any preferred culture of contribution) shows a much more gradual defection (twice as long) to non-cooperation.

We performed multiple simulations to analyze implications of different strategies applied towards population. These simulations included variations with and without members changing strategies over multiple generations. Members changed in continuous distribution modes and optimized their contributing frequency during displacement operations. To model birth, maturation and death of a trait (contributing culture vs. non-contributing) the frequency, the step of displacement, and the deletion events were correlated with the age of the trait (how long the trait has been in place).

**RESULTS AND DISCUSSION**

We propose that differences in the level of patience in a population, as defined above, is a key factor in determining whether a culture of contribution will evolve, or not. Whether people are patient or not may depend on the institutional rules that govern collaboration relative to individual effort. Thus, institutions that incentivize individual effort play a large role in determining whether biologists contribute to community annotations – mainly by affecting the patience of biologists. Our methodology generates three propositions. We then use computer simulations to test these propositions for robustness and validity. We state and discuss our results below.

First of all, we note that the replicator dynamic approach to solving this game allows us to find evolutionary stable strategies. We find that the impatient culture is an

evolutionary stable strategy (ESS) if  $x < \frac{1}{a \frac{1-\delta_H}{1-\delta_L} - 1}$  where  $\delta_H = \frac{p}{1+r_H}$  and  $\delta_L = \frac{p}{1+r_L}$ . This implies that small changes around  $x^* = \frac{1}{a \frac{1-\delta_H}{1-\delta_L} - 1}$  will shift the equilibrium towards a

culture where either contribution prevails or non-contribution prevails. Thus even if, for example, a small  $\epsilon$  proportion of the patient population changes and becomes impatient the culture of the entire population cascades towards the point where only a culture of impatience prevails. The opposite happens if a small  $\epsilon$  proportion of the population mutates and becomes patient. This proposition therefore suggests that the patient culture can prevail only if the proportion  $x$  of patient people in a population satisfies  $x > x^*$ . Since there is no *a priori* reason to believe that  $x$  will always exceed  $x^*$ , we cannot predict that the culture of patience will succeed. For example if  $x^*$  is 8% while the actual proportion of patient biologists in a population is 7.9% then this society of biologists will be doomed to be plagued by non-contribution. But all will be well if the actual proportion of patient biologists happens to be 8.1%. In the context of this paper, this result suggests that a culture that incentivizes non-contribution may well become the dominant culture among biologists even though it is clearly not optimal for the system. Proposition 1 therefore suggests that the populations of biologists will either converge on an equilibrium where everyone is a contributor or where everyone (almost everyone) is a non-contributor because there is no evolutionarily stable equilibrium where both types can coexist. Realistically, both cultures may coexist at any point in time. However, as a dynamic matter one culture will die out. In addition, whether the system reaches the equilibrium where everyone is either a contributor or a non-contributor depends on the initial distribution of contributors relative to non-contributors.

In our computer simulations, biologists have uniform or normally distributed frequencies of contribution. A normal distribution mimics a system where some cultures are more common than others are and around some average frequency

of contribution. A uniform distribution mimics a system where none of the culture is over or underrepresented. Fig. (1) is a representative graph of the replicator dynamic process. We further note here that the biologists in our computer simulation differ only in the frequency of contribution. Recall that we do not explicitly program differences in patience in these simulated biologists. Thus, the results are not an artifact of the simulation program. However, the fitness of the biologist cultures as they interact with each other follows the rank ordering suggested by our theory. Thus, changes in the distribution of biologists in the simulation that matches our theoretical predictions may be evidence in favor of our patience-based theory.

First, we note from Fig. (1) that the simulated population of biologists moves towards an equilibrium with only one culture. This result is sensitive to the choice of initial parameters – the mean of the distribution of the frequency of contribution in the case of the normal distribution or the number of members in each culture in the case of the uniform distribution. This provides evidence in support of our claim that even when the benefits of cooperation by increasing contributions are clear, whether actual cooperation happens or not critically depends on the initial mass of contributors. Thus, formal policies that incentivize contributions (short of actual coercion) might fail if the numbers of biologists who cooperate frequently are not high enough to begin with. We also find that the choice of the normal distribution results in much quicker collapse of the system into one of the extremes relative to the uniform distribution (Fig. 2). Thus, one may conclude that populations with one predominant culture or frequency of contribution are more difficult to sway towards one direction than populations where no culture is predominant. This has important implications for policy implementation. Whether a policy will potentially increase contribution or not depends on the distribution of cooperative biologists in the population.

Of course,  $r_H$ ,  $r_L$ ,  $\alpha$  and  $p$  determine  $x^*$ . This suggests that the ESS are sensitive to changes in patience, the extent to which institutional structures incentivize free riding, and the probability that these institutional structures persist. We therefore ask the following question -- how does  $x^*$  respond to changes in some of the parameters in our model? Recall that  $x^*$  represents a sort of tipping point for whether one culture will prevail over the other or not. Thus if  $x^*$  rises then for the patient culture to succeed a larger proportion of the population needs to be patient. Changes to the actual proportion of a population would be rare and small. We therefore suggest that a rising  $x^*$  makes it harder for the patient culture to succeed. With this outcome in mind we state and discuss our results from propositions 2 and 3 to show how  $x^*$  responds to changes in  $\alpha$  and  $p$ .

Proposition 2 follows from the proof that  $\frac{dx^*}{d\alpha} > 0$ . Clearly the formula for the tipping point in Proposition 1 indicates that as  $\alpha$  rises so does  $x^*$ . This happens because as the incentive to be impatient grows, so does the need for more patient people in society to achieve the efficient outcome. In other words, a rising  $x^*$  implies that a higher proportion of the population needs to be part of a patient culture for a patient culture to prevail. This makes it harder

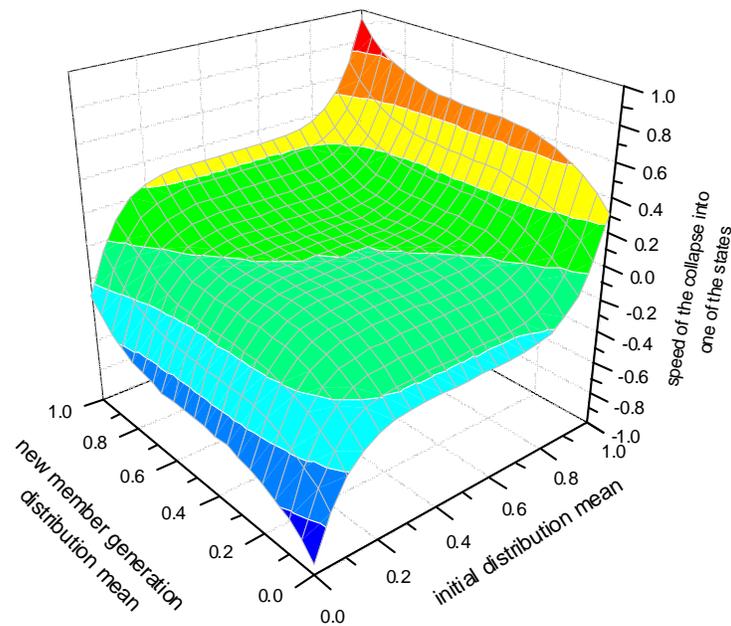
for the patient culture to prevail. For example, say 10% of a society is patient. Further, assume that  $x^*$  is 8%. According to Proposition 1 a patient culture will prevail in this society. Now say that an institutional change incentivizes impatience, e.g., a fall in the contributing standards required by the NSF or National Institutes of Health (NIH) makes  $\alpha$  rise. As a result say  $x^*$  rises to 12%. Now suddenly our society, where the proportion of patient biologists is at 10%, finds itself on the wrong side of  $x^*$  and begins to move towards the impatient equilibrium. In this scenario more and more biologists succumb to non-contribution at the expense of those who do not – effectively crowding out patient biologists. Thus, designing institutional mechanisms to induce biologists to wait are important.

Proposition 3 basically states that as  $p$  falls the present value of future payoffs to both cultures diminish. However, the payoff to the impatient culture falls relatively less than that for the patient culture since the impatient culture places less value on the future anyway. Thus as  $p$  falls the impatient culture becomes fitter (has a relatively higher payoff) than the patient culture. This leads to an increase in the proportion of impatient biologists in the population. Thus, as  $p$  falls, more patient biologists are needed for the patient culture to prevail.

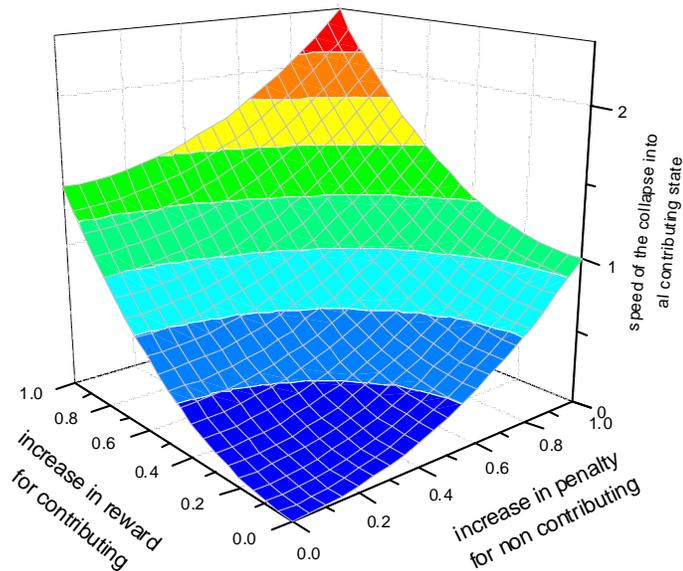
We interpret  $p$  as the probability that a particular institution continues. Thus any change in the rules of the game -- signifying an end to the game itself -- indicates the end of an institution. This proposition then essentially means that if any institutional arrangement becomes more likely to change then it becomes harder for the patient culture to prevail. Say the government is contemplating a change in the regulatory structure that will reduce free riding -- as suggested in Proposition 2 above. Then, paradoxically, even if the purpose of the regulation is to reduce non-contribution and improve efficiency its effect will be to make it harder for biologist's incentives to be patient contributors. This strengthens cultural path dependency – for good or for bad. Moreover, Propositions 2 and 3, taken together suggests that there is a trade off between the expected benefits of any attempt to change the institutional incentive to free ride and the act of the change itself. This sort of trade-off deserves further attention in future research.

Propositions 2 and 3 above show how changes in the parameters that affect the fitness of cultures in a population of biologists affect the likelihood of a culture of contribution taking root. In fact,  $\alpha$ , and  $p$  – the variables of interest in those propositions – affect the rewards and penalties from pursuing a culture of contribution (the patient culture) relative to the rewards and penalties from pursuing a culture of non-contribution (the impatient culture). For example, notice from Table 1 that, an increase in  $\alpha$  rewards an impatient person when she interacts with a patient person. On the other hand, an increase in  $p$  penalizes an impatient person when she interacts with a patient person.

It is clear from equation (10) that the weighted momentum of the distribution plays an important role in the stability of species. Increases in the payoff and penalty ratios essentially shift the momentum of  $c(x)$  thus changing the behavior of the population. This particular treatment of the population of biologists in our simulated environment



**Fig. (2).** The speed by which the population collapses into one of the critical states depends on initial and new member generation distributions. The positive values of the speed signify the convergence to all contributing population and negative values demonstrate the collapse to all defecting population. The actual shape of the saddle depends not only on the means of the distributions but also on standard deviations and matrix parameters describing rewards, penalties, generation change frequencies and lifespan of the members. The figure shows that when selection favors new members for their more contributing traits then the population has a higher chance of eventually moving to the contributing culture (green, yellow, and orange represent increased level of contribution). The shape of the plateau shows a diminishing cultural effect of each generation on the next.

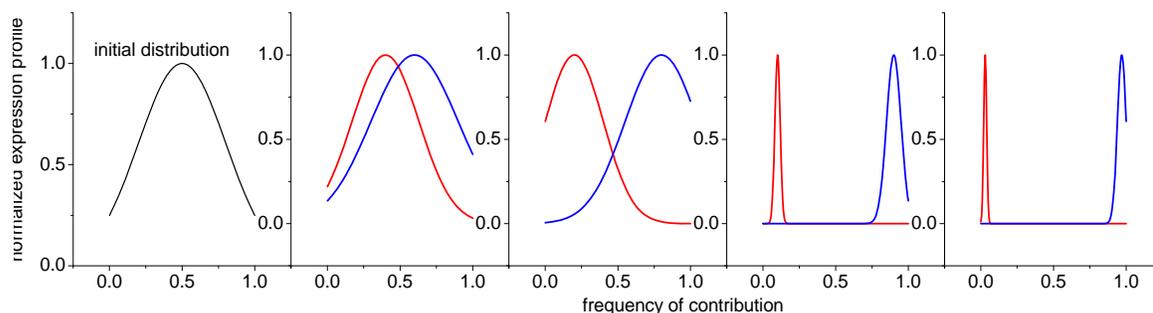


**Fig. (3).** Dependency of the speed by which positive shift in the population occurs from the magnitude of the changes in penalties and rewards. Note that the rewards affect population behavior (contribution) more than the same amount of penalty (compare the left and right corners of the diagram).

captures the effects of Propositions 2 and 3 above – all of which show the impact of changes in the benefits of defection (and therefore the penalties from not contributing) relative to cooperation. Thus, a decrease in  $\alpha$  or an increase in  $p$  will increase the benefit of cooperation relative to the net benefit from defection.

We find, based on reasonable assumptions about initial conditions and fitness (Fig. 3), that an increase in the benefits of contribution, through some reward mechanism,

for example, are more likely to lead to equilibria where the culture of contribution predominates. We find this result is more likely when benefits from contribution rise rather than when penalties from not contributing (through a decrease in the fitness of defection) are increased (Fig. 3 and 4). Our theoretical model predicts that a decrease in  $\alpha$  or an increase in  $p$ , i.e., an increase the benefit of cooperation relative to the net benefit from defection; all make it easier for a culture of patience to take root. Further, given a



**Fig. (4).** Series of expression profile snapshots taken during population dynamics simulations for two different settings of game matrix elements. We model the initial distribution (shown in black on the leftmost diagram) with a non-biased Gaussian with median at the center 0.5 thus giving equal chances for the population to collapse in either one of the states. Comparison of blue series collapsing into all contributing state in contrast with red series collapsing to all defecting to non-contribution clearly demonstrates that an increase in rewarding (CC) by 30% led to significant changes in the future. For this particular simulation the CD has not been changed but decrease in asymmetry of CD $\leftrightarrow$ DC elements has the same effect at a lesser scale. The figure shows that reward (blue line) results in a contributing culture while penalizing (red line) is not enough to push the population to a culture of contribution (please compare this to Fig. 1).

growing culture of patience, people should cooperate more by contributing more. Our simulation directly links an increase in the benefit of cooperation relative to the net benefit from defection to an increase in the number of biologists who contribute more. While we do not program patience in biologists in our simulation, we do order the fitness of the culture they belong to according to our patience-based theory (Table 1). This suggests that our theoretical pathway linking  $\alpha$  and  $p$  to increased contributions by biologists because of their increased desire to cooperate operates through patience. This result is robust to many different specifications for initial conditions and parameters. Moreover, from a policy standpoint this result suggests that funding agencies should reward contributors rather than penalize non-contributors to establish a culture of patience and contribution.

Our theory and simulations suggest that there are no guarantees that a culture of contribution will prevail in a society of biologists – even with the “right” sort of institutions. However, the likelihood that a contributory culture will prevail is enhanced if there is an increase in the rewards from contribution relative to the penalty from not contributing. In particular we point toward the three lynchpins of any policy to make a culture of contribution the predominant culture in this society – the stability of rules, the patience of individuals and the rewards from contributing relative to not contributing.

Our model and simulations have implications for experimental design as well. In cooperation experiments with human subjects patience is usually modeled using experiment termination rules [25]. In these experiments, games have some probability of ending. Mathematically, of course, the probability a game continues is equivalent to a discount rate. Thus, using game termination rules to model discount rates may be mathematically justified. Further, presumably, if the subjects know that a game is very likely to end they will behave more impatiently than if they knew that the game is more likely to continue. But this psychological justification is muddy at best. If people come into the experiment with a culturally predetermined patience level then it is unclear what sort of behavior is being captured in an experiment where the patience level is forced on the

subject as *a part of the experimental treatment*. Thus, if one accepts that patience is culturally determined, as appears to be the broad consensus among behavioral economists and psychologists [18, 21], then experiments that force specific patience levels on subjects may be capturing behavior whose interpretation is at best unclear [27]. We have developed a framework that allows patience and cooperation to co-evolve in an experimental setting. When patience is not allowed to evolve by experimental design neither can cooperation. This may be why cooperation is so hard to replicate in the lab [25, 26]. Thus, for example, in order to get human subjects to cooperate in experimental settings all subjects should be faced with the same experiment termination rules. Further, some sort of mechanism could sort subjects according to their patience levels prior to the experiment. Now one could test if patience evolves in settings where our theory and simulations suggests it should. If patience evolves so should cooperation according to the Folk Theorem. After all biology and human life is replete with examples of cooperation [37]! Our theory provides the levers to engineer this cooperation in the lab and, if experimentally justified, in the real world.

## CONCLUSION

Our theory and simulations suggest the there are no guarantees that a culture of contribution will prevail in a society of biologists – even with the “right” sort of institutions. However, a culture of contribution is more likely if contributors are rewarded rather than by punishing non-contributors. This can be done by manipulating the stability of rules, the patience of individuals, and the rewards from contributing relative to not contributing. Thus our theory suggests several testable hypotheses and is a first attempt at developing a systematic methodology for organizing scientists to maximize scientific cooperation.

## CONFLICT OF INTEREST

The authors confirm that this article content has no conflicts of interest.

## ACKNOWLEDGEMENTS

A part of this work has been developed with the help of funding from the NSF (Grant number 904674, PI Atin

Basuchoudhary). We would like to thank Dr. Srimayi Sen and Dr. Peter Grajzl for numerous comments that have improved our paper. Any deficiencies are entirely ours. This work stems from the feedback to Raja Mazumder during the panel discussions on Outreach to User Communities & Community Annotation in the 4<sup>th</sup> International Biocuration Meeting, Tokyo, Japan, Oct. 11 – 14, 2010.

## REFERENCES

- [1] J. Vukov, F. L. Pinheiro, F. C. Santos, and J. M. Pacheco, "Reward from punishment does not emerge at all costs," *PLoS Comput. Biol.*, vol. 9, p. e1002868, 2013.
- [2] R. Mazumder, D. A. Natale, J. A. Julio, L. S. Yeh, and C. H. Wu, "Community annotation in biology," *Biol Direct*, vol. 5, p. 12, 2010.
- [3] T. Clark and J. Kinoshita, "Alzforum and SWAN: the present and future of scientific web communities," *Brief Bioinform*, vol. 8, pp. 163-171, 2007.
- [4] S. L. Salzberg, "Genome re-annotation: a wiki solution?," *Genome Biol.*, vol. 8, p. 102, 2007.
- [5] K. Wang, "Gene-function wiki would let biologists pool worldwide resources," *Nature*, vol. 439, p. 534, 2006.
- [6] J. Giles, "Key biology databases go wiki," *Nature*, vol. 445, p. 691, 2007.
- [7] J. Thornton, "Annotations for all by all - the BioSapiens network," *Genome Biol.*, vol. 10, p. 401, 2009.
- [8] B. Mons, M. Ashburner, C. Chichester, E. van Mulligen, M. Weeber, J. den Dunnen, G. J. van Ommen, M. Musen, M. Cockerill, H. Hermjakob, A. Mons, A. Packer, R. Pacheco, S. Lewis, A. Berkeley, W. Melton, N. Barris, J. Wales, G. Meijssen, E. Moeller, P. J. Roes, K. Borner, and A. Bairoch, "Calling on a million minds for community annotation in WikiProteins," *Genome Biol.*, vol. 9, p. R89, 2008.
- [9] R. Hoffmann, "A wiki for the life sciences where authorship matters," *Nat. Genet.*, vol. 40, pp. 1047-1051, 2008.
- [10] H. Stehr, J. M. Duarte, M. Lappe, J. Bhak, and D. M. Bolser, "PDBWiki: added value through community annotation of the Protein Data Bank," *Database (Oxford)*, vol. 2010, p. baq009, 2010.
- [11] J. Daub, P. P. Gardner, J. Tate, D. Ramskold, M. Manske, W. G. Scott, Z. Weinberg, S. Griffiths-Jones, and A. Bateman, "The RNA WikiProject: community annotation of RNA families," *RNA*, vol. 14, pp. 2462-2464, 2008.
- [12] J. E. Loveland, J. G. Gilbert, E. Griffiths, and J. L. Harrow, "Community gene annotation in practice," *Database (Oxford)*, vol. 2012, p. bas009, 2012.
- [13] V. Guignon, G. Droc, M. Alaux, F. C. Baurens, O. Garsmeur, C. Poirion, T. Carver, M. Rouard, and S. Bocs, "Chado controller: advanced annotation management with a community annotation system," *Bioinformatics*, vol. 28, pp. 1054-1056, 2012.
- [14] B. M. Good, E. L. Clarke, L. de Alfaro, and A. I. Su, "The Gene Wiki in 2011: community intelligence applied to human gene annotation," *Nucleic Acids Res.*, vol. 40, pp. D1255-1261, 2012.
- [15] P. P. Gardner, J. Daub, J. Tate, B. L. Moore, I. H. Osuch, S. Griffiths-Jones, R. D. Finn, E. P. Nawrocki, D. L. Kolbe, S. R. Eddy, and A. Bateman, "Rfam: Wikipedia, clans and the "decimal" release," *Nucleic Acids Res.*, vol. 39, pp. D141-145, 2011.
- [16] G. M. Dall'Olio, J. Bertranpetit, and H. Laayouni, "The annotation and the usage of scientific databases could be improved with public issue tracker software," *Database (Oxford)*, p. baq035, 2010.
- [17] R. D. Finn, P. P. Gardner, and A. Bateman, "Making your database available through Wikipedia: the pros and cons," *Nucleic Acids Res.*, vol. 40, pp. D9-12, 2012.
- [18] O. S. Curry, M. E. Price, and J. G. Price, *Patience is a virtue: cooperative people have lower discount rates. Personal Individual Differences*, vol. 44, pp. 778-783, 2008.
- [19] D. Fudenberg and E. Maskin, "The folk theorem In repeated games with discounting or with incomplete information.," *Econometrica*, vol. 54, pp. 533-554, 1986.
- [20] Y. Lengwiler, "Heterogeneous patience and the term structure of real interest rates.," *Am. Econ. Rev.*, pp. 890-896, 2005.
- [21] N. Wilkinson, "An introduction to behavioral economics." New Palgrave Macmillan, York, NY, pp. 189-263, 2008.
- [22] J. T. Warner and S. Pleeter, "The personal discount rate: Evidence from military downsizing programs.," *Am. Econ. Rev.*, vol. 91, pp. 33-53, 2001.
- [23] G. Clark, *A farewell to alms*. Princeton University Press: Princeton, NJ, 2007.
- [24] A. Basuchoudhary, T. Siemers, and A. Sam, "Civilization and the evolution of short sighted agents," *Va Econ. J.*, pp. 11-29, 2010.
- [25] P. Dal-Bo and R. F. Guillaume, "The Evolution of cooperation in infinitely repeated games: Experimental evidence. ," *Am. Econ. Rev.*, vol. 101, pp. 411-429, 2011.
- [26] J. Duffy and J. Ochs, "Cooperative behavior and the frequency of Social interaction.," *Games Econ. Behav.*, vol. 66, pp. 785-812, 2009.
- [27] H. T. Normann and B. Wallace, *The impact of the termination rule on cooperation in a prisoner's dilemma experiment. Int. J. Game Theory*, vol. 41, pp. 707-718, 2011.
- [28] E. Boutet, D. Lieberherr, M. Tognolli, M. Schneider, and A. Bairoch, "UniProtKB/Swiss-Prot," *Methods Mol. Biol.*, vol. 406, pp. 89-112, 2007.
- [29] A. Bairoch, B. Boeckmann, S. Ferro, and E. Gasteiger, "Swiss-Prot: juggling between evolution and stability," *Brief Bioinform*, vol. 5, pp. 39-55, Mar 2004.
- [30] J. M. Smith, *Evolution and the theory of games*. Cambridge University Press, Cambridge 1982.
- [31] H. Gintis, *Game theory evolving*. Princeton University Press, Princeton 2000.
- [32] T. Sandler, "Collective Action and Transnational Terrorism," *World Econ.*, vol. 26, pp. 779-802, 2003.
- [33] J. Bendor and P. Swistak, "The evolutionary stability of cooperation.," *Am. Polit. Sci. Rev.*, vol. 91, pp. 290-307, 1997.
- [34] M. Milinski, "Tit for tat in sticklebacks and the evolution of cooperation," *Nature*, vol. 325, pp. 433-435, 1987.
- [35] G. J. Mailath, L. Samuelson, and A. Shaked, "Endogenous interactions.," In: *CARESS Working Papres endo-one*, University of Pennsylvania Center for Analytic Research and Economics in the Social Sciences, USA 1997.
- [36] J. E. Harrington, *Games, strategies and decision making*. Worth, New York: 2009.
- [37] M. Nowak and R. Highfield, *SuperCooperators: Altruism, evolution, and why we need each other to succeed*. New York: A division of Simon and Schuster, Inc., Free Press 2011.

Received: February 11, 2013

Revised: May 03, 2013

Accepted: May 08, 2013

© Basuchoudhary et al.; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.