Variants at serotonin transporter and 2A receptor genes predict cooperative behavior differentially according to presence of punishment

Kari B. Schroeder[a,b,1], Richard McElreath[b,c], and Daniel Nettle[a]

[a]Centre for Behavior and Evolution, Institute of Neuroscience, Newcastle University, Newcastle Upon Tyne NE2 4HH, United Kingdom; and [b]Department of Anthropology and [c]Center for Population Biology, University of California, Davis, CA 95616

Edited by Robert M. May, University of Oxford, Oxford, United Kingdom, and approved January 17, 2013 (received for review September 27, 2012)

Punishment of free-riding has been implicated in the evolution of cooperation in humans, and yet mechanisms for punishment avoidance remain largely uninvestigated. Individual variation in these mechanisms may stem from variation in the serotonergic system, which modulates processing of aversive stimuli. Functional serotonin gene variants have been associated with variation in the processing of aversive stimuli and widely studied as risk factors for psychiatric disorders. We show that variants at the serotonin transporter gene (*SLC6A4*) and serotonin 2A receptor gene (*HTR2A*) predict contributions to the public good in economic games, dependent upon whether contribution behavior can be punished. Participants with a variant at the serotonin transporter gene contribute more, leading to group-level differences in cooperation, but this effect dissipates in the presence of punishment. When contribution behavior can be punished, those with a variant at the serotonin 2A receptor gene contribute more than those without it. This variant also predicts a more stressful experience of the games. The diversity of institutions (including norms) that govern cooperation and punishment may create selective pressures for punishment avoidance that change rapidly across time and space. Variant-specific epigenetic regulation of these genes, as well as population-level variation in the frequencies of these variants, may facilitate adaptation to local norms of cooperation and punishment.

public goods game | collective action | behavioral plasticity | 5-HTTLPR

Punishment has likely been a strong selective force in human evolutionary history. The punishment of free-riders enables cooperation (1), which is a hallmark of human evolution. Across diverse cultures, social norms, both within and outside of the domain of cooperation, are enforced with punishment ranging from gossip to exile and even death (2). Thus, natural selection should shape cognitive and affective mechanisms that enable the internalization of norms (3), sensitivity to the probability of punishment by others for norm violation, and aversion to imagined or experienced punishment (4).

However, mechanisms for punishment avoidance remain largely uninvestigated. Variation in the serotonergic system could underlie individual variation in psychological mechanisms for avoiding punishment. Prediction of (5) and response to (6) negative outcomes and social decision-making behavior (7) can be modified via manipulation of serotonin levels. A bias toward negative stimuli (8) characterizes mood and anxiety disorders; altered regulation of the serotonergic system has long been implicated in these disorders. Processing of aversive stimuli and sensitivity to the social environment are also linked to functional serotonin gene variants. A length polymorphism (5-HTTLPR) in the promoter region of the serotonin transporter gene, *SLC6A4* [*solute carrier family 6 (neurotransmitter transporter, serotonin), member 4*, also referred to as 5-HTT], predicts increased observational fear conditioning (9) and amygdala activation in the presence of threatening social cues (10). This polymorphism, as well as a polymorphism in the promoter region of the serotonin 2A receptor gene, *HTR2A* [*5-hydroxytryptamine (serotonin) receptor 2A, G protein-coupled*, also referred to as 5-HT_{2A}], is also associated with the personality dimension of neuroticism (11, 12), increased risk for depression and other psychiatric disorders (13, 14), and increased cortisol response to a psychosocial stressor (12, 15). Intriguingly, a recent assessment of global variation at *SLC6A4* and *HTR2A* suggests unusual evolutionary histories at these loci in humans (16). Haplotypes (i.e., the combination of linked alleles that are inherited as a unit) at these loci have a striking geographic distribution, may have been under directional selection, and are estimated to have originated or spread relatively recently (16).

We hypothesized that individuals with particular variants at *SLC6A4* and *HTR2A* are more sensitive to punishment and will thus be more cooperative than individuals without these variants when noncooperative behavior can be punished. We used a standard of experimental economics, the Public Goods Game (PGG), to investigate the effect of *SLC6A4* and *HTR2A* haplotypes on cooperative behavior in the presence and absence of punishment and assessed sensitivity to punishment via changes in affect and cortisol secretion during the PGG. In the PGG, each player privately decides how much of her money to contribute to the public good, the total of which is multiplied by a number greater than one and divided equally among players. Although everyone in the group benefits equally from contributions, an individual maximizes her payoff by keeping her money. The PGG has been used to study how contributions change when players are given the opportunity to punish each other (via fines). In the absence of punishment opportunities, contributions decline over rounds. Punishment targeted at low contributors attenuates that decline (17, 18).

One hundred eighty-four students participated in the study at Newcastle University. Participants remained in groups of four for the duration of the experiment and played 10 rounds each of two versions of the PGG. In the No Punishment game, each player received 20 tokens per round and privately decided how many tokens (integer from 0 to 20) to contribute to the group fund. The Punishment game always followed the No Punishment game. It differed in that after players' contributions and incomes for a given round were revealed, players assigned 0 to 10 negative tokens to each other player. Each negative token cost the giver one token and the recipient three tokens. Participants were paid at the end of the experiment (one token: £0.015).

Before the games, self-reported assessments of personality and depression were collected. Self-reported positive and negative affect was also assessed at five times during the experiment. DNA was extracted from buccal swabs collected at the start of

Author contributions: K.B.S. conceived of the study; K.B.S., R.M., and D.N. designed the experiment; K.B.S. performed research; K.B.S. and R.M. contributed new reagents/analytic tools; K.B.S., R.M., and D.N. analyzed data; and K.B.S. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. E-mail: kbschroeder@ucdavis.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1216841110/-/DCSupplemental.

Table 1. Fixed-effects coefficients and variance components for the best models of the number of tokens contributed to the group fund

	Model 1			Model 2			Model 3		
Parameter	Estimate	2.5%	97.5%	Estimate	2.5%	97.5%	Estimate	2.5%	97.5%
Fixed effects									
Intercept	−0.02 (0.36)	−0.7	0.674	1.01 (0.45)	0.12	1.900	0.01 (0.54)	−1.046	1.075
SH1	1.33 (0.37)	0.6	2.058				1.38 (0.39)	0.619	2.132
SH1 × Lag MCO	−0.08 (0.03)	−0.1	−0.030				−0.09 (0.03)	−0.143	−0.034
SH1 × Lag punished	−0.22 (0.08)	−0.4	−0.073				−0.23 (0.08)	−0.382	−0.081
HH1				−0.17 (0.44)	−1.03	0.700	0.02 (0.45)	−0.855	0.897
HH1 × P game				1.11 (0.36)	0.40	1.818	1.01 (0.36)	0.298	1.719
P game	1.44 (0.17)	1.1	1.769	0.58 (0.34)	−0.08	1.241	0.66 (0.34)	−0.004	1.316
Round	−0.05 (0.03)	−0.1	0.004	−0.06 (0.03)	−0.11	−0.001	−0.06 (0.03)	−0.113	−0.003
First round	7.71 (0.31)	7.1	8.325	7.69 (0.32)	7.07	8.306	7.61 (0.32)	6.995	8.232
Lag contribution	0.33 (0.02)	0.3	0.359	0.32 (0.02)	0.29	0.355	0.32 (0.02)	0.286	0.355
Lag MCO	0.56 (0.03)	0.5	0.612	0.50 (0.02)	0.46	0.543	0.56 (0.03)	0.502	0.613
Lag punished	0.24 (0.06)	0.1	0.366	0.10 (0.04)	0.02	0.177	0.25 (0.07)	0.122	0.379
Variance components									
Participant	3.14 (1.77)			3.23 (1.80)			3.27 (1.81)		
Residual	16.21 (4.03)			16.10 (4.01)			16.04 (4.01)		

Predictions from models 1–3 are plotted in Figs. 1 and 2 and Fig. S1, respectively. Parentheses contain SEs or, for the variance components, SDs of the estimates. Lag refers to the previous round. *First round*, the initial round of either game (i.e., round 1 or round 11); *Lag contribution*, the lagged contribution of ego; *Lag MCO*, the lagged mean contribution of the group, excluding ego; *Lag punished*, the lagged number of negative tokens ego received; *P game*, the punishment game.

the experiment. Two variants in each gene were genotyped, and haplotypes for *SLC6A4* and *HTR2A* were inferred for 177 and 174 participants, respectively. Haplotypes were classified as *SLC6A4* 1 (*SH1*) or *SLC6A4* 2 (*SH2*) and *HTR2A* 1 (*HH1*) or *HTR2A* 2 (*HH2*) (Table S1 and *SI Text: Haplotype Classification*).

To investigate the effects of the haplotypes on the number of tokens contributed to the group fund, we constructed a base regression model with game variables that previous studies have shown to be important predictors of contributions (17, 19) (Table 1) and varying intercepts for individuals. Analyzing each gene separately, we then introduced haplotype into the models, investigating different relationships between haplotype and phenotype as well as interactions between haplotype and the game variables. Model selection was conducted via Akaike Information Criterion (AIC) (20). Results from the best model for each gene, as well as the model that combines the best model for each gene, are presented in Table 1.

Results and Discussion

The effect of *SLC6A4* and *HTR2A* haplotypes on contribution behavior depends on the game played. *SH2* homozygotes (33.90% of participants) contributed less in the No Punishment game (Fig. 1, Table 1, Fig. S1, and Table S2). The predicted contribution for *SH2* homozygotes in round 1 is 7.65 (7.05, 8.31) tokens, compared with 8.96 (8.41, 9.43) tokens for *SH1* homozygotes and heterozygotes. (Parentheses contain 95% confidence intervals for predictions.) By the final round of the No Punishment game, predicted contributions are 1.94 (1.50, 2.40) and 4.90 (4.43, 5.30) tokens for *SH2* homozygotes and those with *SH1*, respectively. In the Punishment game, this effect was diminished. The presence of punishment opportunities stemmed the decay in contributions of *SH2* homozygotes (Fig. 1, Table 1, Fig. S1, and Table S2). This game-dependent behavior is consistent with the interpretation that *SH1* homozygotes and heterozygotes experience greater norm internalization. It is also concordant with the explanation that they are more averse to harming others, an outcome that has been experimentally influenced via manipulation of serotonin levels (7).

In the No Punishment game, the contribution behavior of *HH1* homozygotes and heterozygotes (82.76% of participants) is not discernible from that of *HH2* homozygotes (Fig. 2, Table 1, Fig. S1, and Table S2). However, in the Punishment game, participants with *HH1* contributed more than *HH2* homozygotes. For the final round of the Punishment game, the predicted contribution for those with *HH1* is 12.31 (11.86, 12.78) tokens, compared with 9.47 (8.73, 10.26) tokens for *HH2* homozygotes. The higher contributions in the Punishment game for participants with *HH1* did not depend upon the number of negative tokens received by participants in the previous round. The mere introduction of explicit punishment opportunities induced higher contributions in those

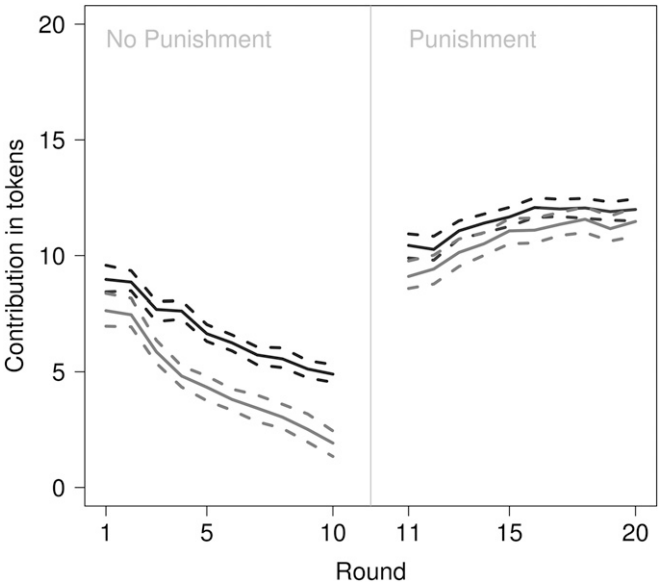

Fig. 1. Predicted contributions from model 1. *SH1* homozygotes and heterozygotes are dark gray and *SH2* homozygotes are light gray. Dotted lines illustrate 95% confidence intervals.

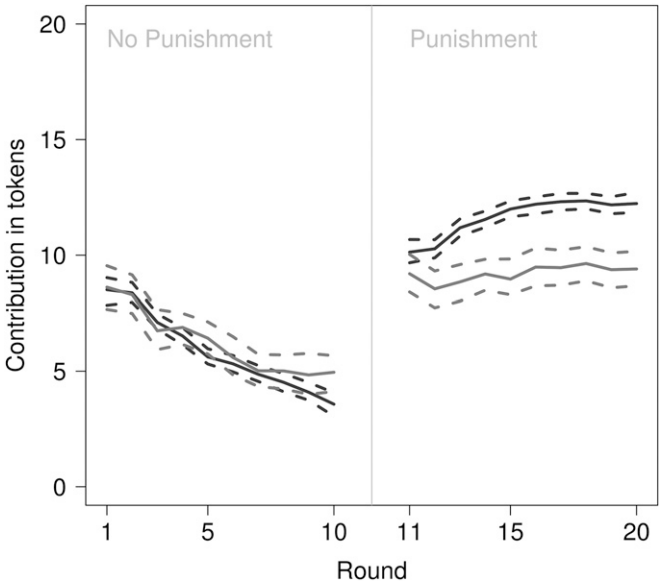

Fig. 2. Predicted contributions from model 2. *HH1* homozygotes and heterozygotes are dark gray and *HH2* homozygotes are light gray. Dotted lines illustrate 95% confidence intervals.

with *HH1* relative to those homozygous for *HH2*. This is in agreement with the interpretation that individuals with *HH1* are more averse to imagined punishment or have a higher expectation of being punished. Because the Punishment game always followed the No Punishment game, the role of serotonin and the 2A receptor in reversal learning (21) is also relevant.

In our experiment, groups were formed without prior knowledge of genetic variation. However, there are still detectable effects of the groups' haplotypic composition on contributions. Groups with three or four participants with one or two copies of *SH1* had substantially higher mean contributions compared with groups with zero to two participants with *SH1* (Fig. 3). This group-level difference was erased in the Punishment game (Fig. 3). This result complements previous work that has demonstrated the possibility for group-level outcomes to be influenced by individual variation in cooperative behavior (22, 23).

We do not observe an association between our measures of neuroticism or depression and variation at *SLC6A4* or *HTR2A*. However, individuals homozygous for *HH1* (27.01% of participants) felt worse as a result of participating in the games (Fig. S2 and Table S3). Negative affect (NA) for participants with two copies of *HH1* began to increase relative to that for other participants after the introduction of the No Punishment PGG. Predicted NA by the end of the experiment is 8.04 (7.16, 9.02) for those homozygous for *HH1* and 7.01 (6.50, 7.52) for those with one or two copies of *HH2*. Predicted NA is not, however, higher for *HH1* homozygotes before the games (Fig. S2 and Table S3). A relationship between *HH1* and a more stressful experience of the games is also indicated by higher cortisol secretion during the Punishment game for nondepressed females with *HH1* (Table S4). Predicted cortisol secretion during the Punishment game for females in the lowest quartile of the sex-specific distribution of depression scores is 140.87 nmol/L (114.11, 172.29 nmol/L) for females with *HH1*, compared with 56.74 nmol/L (34.96, 93.06 nmol/L) for females without *HH1*. Although mild, these effects demonstrate the potential psychological cost of an aversion to or expectation of punishment and are relevant to the hypothesized complex interaction of the serotonergic system, hypothalamic pituitary adrenal axis, and stress exposure in the development of depression (24, 25).

A defining characteristic of human evolutionary history is the diversity of institutions and norms that shape cooperation and punishment (18, 26). Thus, selective pressures for punishment avoidance may vary with cultural environments. For example, in a corrupt society, the probability of being punished for a violation of a law may be unpredictable and depend little on the actor's behavior. Such an outcome is similar to that observed by (18). Herrman et al. (18) demonstrated that in countries with a weak rule of law, punishment in the PGG is not strongly biased toward those who give less than the punisher, unlike in countries with a strong rule of law, in which punishment is heavily biased toward those who give less than the punisher. When punishment is highly unpredictable, the evolutionary costs and benefits of a psychology that is more averse to punishment may be altered.

However, mechanisms upon which selection may act to drive behavioral adaptation to local norms of cooperation and punishment remain largely unknown. Our results suggest that the effect of *SLC6A4* and *HTR2A* variation on cooperative behavior may vary depending upon aspects of the social context, including opportunities for behavior to be punished. Thus, substantial population-level variation in frequencies of *SH1* and *HH1*, as well as evidence of potential selection at *SLC6A4* and *HTR2A* and a very recent estimated age (19,000 ± 4,000 y ago) for *SH1* (16) are provocative. (Here, we refer to a subset of *SH1* further characterized by the derived allele at reference single nucleotide polymorphism 1042173 (rs1042173). See *SI Text*: *Robustness of Contribution Inferences*: *Characterization of haplotypes* and ref. 16.)

Of equal interest is recent molecular evidence for genotype- or haplotype-specific epigenetic regulation at *HTR2A* (27, 28) and *SLC6A4* (29, 30). These results, in concert with those from gene by environment studies (31, 32), lend increasing support to the hypothesis that, rather than conferring susceptibility to psychopathology, polymorphisms at *SLC6A4* and *HTR2A* enable increased plasticity of the serotonergic system in response to the social environment (31, 32) (*SI Text: Haplotype Classification*). This precludes a universal assumption of how *SH1*, *SH2*, *HH1*, and *HH2* affect serotonergic functioning. Moreover, it suggests an additional route by which cross-cultural variation in

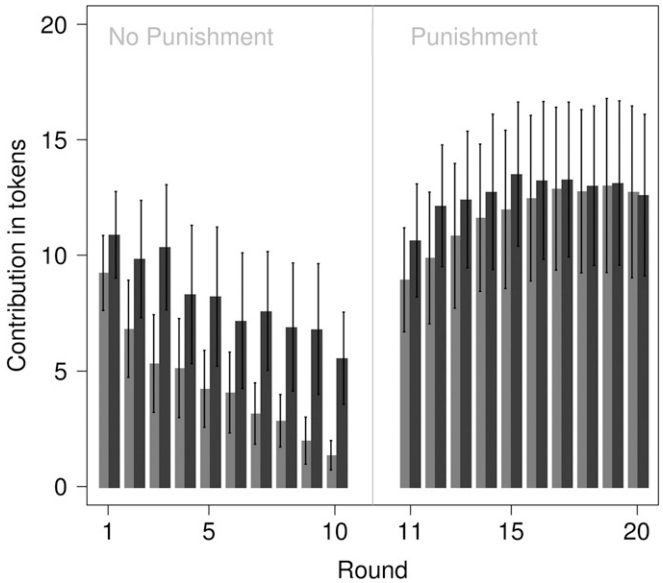

Fig. 3. Mean and 95% confidence intervals of mean contributions for groups characterized by the number of participants with *SH1*. Groups with three or four participants with *SH1* are dark gray, and those with zero, one, or two participants with *SH1* are light gray.

cooperation could be produced. Whereas the ability of cultural environments to shape selective pressures has received attention (33, 34), and has even been implicated in the global distribution of the short allele at 5-HTTLPR (35), the possible role of epigenetic regulation in this scenario has not. Genetic variation that enables behavioral plasticity could facilitate the rapid evolution of behavior in response to changing environments (36), including sanctioning institutions and other aspects of the sociocultural environment.

Methods

Experimental Sessions. Eleven sessions were conducted from November 2010 to March 2011 in computer clusters on the Newcastle University campus. The number of participants per session ranged from eight to 28. Participants were spaced such that there was either an empty computer or wall immediately adjacent to both sides of each participant. They were instructed not to communicate with each other in any way, including eye contact or body language. A purpose-built website was used to communicate all instructions to participants, administer questionnaires, and conduct the PGG.

Participants. One hundred and eighty four participants (77 males; mean age, 20.8 y) were recruited through the university psychology student mailing list, advertisements on the Web site of Newcastle University, a participant pool maintained by the Newcastle University Institute of Neuroscience, and flyers posted on campus. The study received approval from the Newcastle University Medical School Board of Ethics before commencement. All participants gave their written consent to participate in the study. Participation criteria included the following: fluency in written English, a minimum age of 18 y, and no psychiatric or steroid medications. Subjects received either a show-up fee or course credit (the latter option for psychology students only). A show-up fee of £3 was increased to £5 for the last six sessions to motivate participation.

Public Goods Game. The PGG structure used closely follows that of ref. 18. After reading instructions for the No Punishment game, participants had to correctly answer a set of questions designed to assess their understanding before proceeding. Participants were told only that they would be introduced to a different version of the game after playing the current game for 10 rounds. The marginal per capita return on the public good (the sum of tokens contributed by all group members to the project) was set at 0.4 tokens. Following the contribution stage of each round, each player was shown the contribution and income of all players in her group. Cumulative income for the game was summarized at the end of each round. Player identity could not be tracked from round to round.

Participants were then introduced to the Punishment game. They had to correctly answer questions designed to assess their understanding of the new version of the game before proceeding to 10 rounds of the Punishment game. After assigning 0 to 10 negative tokens to each player in a given round, each player saw a summary screen that included the number and cost of negative tokens given and received and income adjusted for the cost of negative tokens. Participants were immediately paid their earnings and show-up fee in cash after completing the experiment.

Genotyping. DNA was extracted from buccal samples at Newcastle University with the Isohelix DNA Isolation kit (trademarked product of Cell Projects Ltd.). Samples were genotyped by NewGene for the length polymorphism in the promoter region (5-HTTLPR) and variable number of tandem repeats in intron 2 (serotonin transporter intron 2 variable number of tandem repeats, STin2 VNTR) of *SLC6A4* as well as two single nucleotide polymorphisms (SNPs) in *HTR2A*, rs6311, and rs6313 (positions −1438 in the promoter and 102 in exon 1, respectively).

Cortisol. Participants were instructed to refrain from strenuous exercise and alcohol the day of the experiment and from having a meal or caffeine within two hours of the start of the experiment. Sessions all commenced at 1430 hours. Saliva was collected with the Sarstedt Salivette at three time points during the experiment: the beginning of the experiment (T1), 15 min after the end of the No Punishment game (T2), and 15 min after the end of the Punishment game (T3). For the first 8 min of each 15-min waiting period after the No Punishment and Punishment games, participants completed a self-assessment of positive affect (PA) and NA and then watched nature videos. The remaining final seven minutes of each waiting period were spent either reading the instructions to the Punishment game, answering questions that tested understanding of the game, and completing an additional self-assessment of positive and negative affect (T2) or being debriefed about the experiment (T3). Cortisol was assayed in duplicate for each sample at the laboratory of C. Kirschbaum (University of Dresden, Dresden, Germany). The average of each pair of measurements was calculated and used as the measurement for that sample. Salivary cortisol levels generally decreased over the duration of the experiment, possibly attributable to time of day (all experiments started at 1430 hours).

Self-Reported Assessments and Background Information. Before the PGG, participants completed a self-reported personality assessment, major depression inventory, and baseline assessment of PA and NA. Personality was assessed with a 120-item version of the International Personality Item Pool version of the NEO-PI-R.* Depression was assessed with the 10-item Major Depression Inventory (MDI) (37). PA and NA were assessed at five different times with the 10-item International Positive and Negative Affect Schedule short form (38): PA1 and NA1, beginning of the experiment; PA2 and NA2, after reading the PGG instructions (No Punishment version) and before commencing the game; PA3 and NA3, immediately after the No Punishment version of the PGG; PA4 and NA4, after reading instructions and before commencing the Punishment game; and PA5 and NA5, immediately after the Punishment game. The following information was also collected for each participant: age, sex, use of hormonal contraceptives, and "biological ancestry" (options were: Subsaharan African, Northern African, Southern European, Northern European, Eastern European, West Asian, Central Asian, East Asian, Southeast Asian, multiple origins, and no response).

Haplotype Phasing and Classifications. Haplotype phase was estimated separately for each locus with the software PHASE Version 2.1.1 (39–41). The only imputed genotype included was for one individual at rs6313. The most likely haplotype pairs for each individual as estimated by PHASE were used in downstream analyses (Table S1). Our schema for grouping haplotypes were similar to that of ref. 16, which classified haplotypes at *HTR2A* and *SLC6A4* as predicted high or low expression based on published molecular studies. Our schema differs from ref. 16 in that we do not use data on rs6312 (position −783 relative to the start of transcription) in *HTR2A*. Also, because of increasing evidence that the effect of these variants on expression is under epigenetic regulation (*SI Text: Haplotype Classification*), we note that "differential expression" may be a more appropriate description than high or low expression. *SH1* is characterized by the short allele at 5-HTTLPR and the 12-repeat allele at STin2 VNTR. *HH1* is characterized by rs6311G and rs6313C alleles (referred to in ref. 16 as −1438G and 102C). We investigate the robustness of our results to the haplotype classifications used (Fig. S3 and *SI Text: Robustness of Contribution Inferences: Characterization of haplotypes*).

Data Analysis. We analyzed the data and created all figures in the R statistical and computing environment (42–47).

Contributions. To predict contributions, we constructed a base model that includes the following game variables: *P game* (binary; whether Punishment game), *Round*, *First round* (first round of either game), *Lag contribution* (lagged contribution of ego; lagged refers to the previous round), *Lag MCO* (lagged mean contribution of group members, excluding ego), and *Lag punished* (lagged number of negative tokens received). These predictors are consistent with important predictors of contributions from previous studies (17, 19). More complex models that include interactions among these game variables were not used as they resulted in little change in predicted contributions. Random intercepts for groups were also considered but not included. For the Gaussian model, the among-group variance estimate decreases from 17.77 ($\sigma = 4.22$) to 0.39 ($\sigma = 0.63$) when the six game variables above are included in the base model. This largely results from the inclusion of *Lag contribution* and *Lag MCO*. Predicted contributions from this base model are consistent with classic PGG outcomes (e.g., ref. 17) (Fig. S1).

In assessing the effect of the *SLC6A4* and *HTR2A* on contributions, we considered three possible relationships between *SH1* or *HH1* and phenotype: dominant (Hd), recessive (Hr), and incomplete dominant (Hi). For each possible phenotype, we iterated over the base model, first including phenotype as a main effect and then interacting it with each of the control variables. Models with AIC (20) weights greater than 0.05 were combined until more

*Johnson JA, *Development of a short form of the IPIP-NEO Personality Inventory*. Second Biennial Meeting of the Association for Research in Personality, June 16–18, 2011, Riverside, CA.

complex models were not supported. The best models for each phenotype were then compared.

For *SLC6A4* Hd and Hi, the best models each include an interaction between phenotype and *Lag MCO*, as well as phenotype and *Lag punished*. The best Hd model receives far greater support than the best Hi and Hr models (predictions plotted in Fig. 1). For *HTR2A*, the best Hd and Hi models for each include an interaction between phenotype and *P game*. The *HTR2*A Hd model receives most of the support (predictions plotted in Fig. 2), although there is some support for Hi and Hr as well. The *SLC6A4* and *HTR2A* models that received the most support were then combined. Individuals with one or two copies of *SH1* are not more or less likely to have one or two copies of *HH1* ($\chi^2 = 2.29$; df = 1; $P = 0.13$). Predictions from the combined model, which receives far more support than either model alone, are plotted in Fig. S1, and the coefficients are presented in Table 1 (model 3). Henceforth, we refer to this model as the "best candidate model."

Predicted contributions were generated from samples from the posterior density of the model, assuming a multivariate normal density. This was done 100 times for each combination of round, game, and haplotype. The mean and 2.5 and 97.5 percentiles of the predicted values are plotted.

Robustness of the results to different model families, ethnic composition, sex, and haplotype classification was confirmed (Figs. S1, S3, and S4; Tables S2 and S5; and *SI Text: Robustness of Contribution Inferences*).

Neuroticism and depression. Regression analyses were used to separately assess the effect of *SLC6A4* and *HTR2A* haplotypes on neuroticism. Three specifications of the relationship between haplotype and phenotype were used for each locus: dominant, recessive, and incomplete dominant. AIC (20) was then used to select among these models for each locus. For each locus, models without any genetic information receive the least support, and models specifying a dominant relationship between *SHI* or *HH1* and phenotype receive the most support. However, the 95% confidence intervals for the coefficients for SHI and *HH1* overlap substantially with zero.

An effect of *SLC6A4* and *HTR2A* haplotypes on MDI (square root of Major Depression Inventory Score) was assessed in the same manner as neuroticism. Results were similar in that the least favored models are those that do not include any genetic information, and the models that receive the most support specify a dominant relationship between *SH1*/*HH1* for each locus and phenotype. However, as with neuroticism, the 95% confidence intervals for the coefficients for *SH1* and *HH1* overlap extensively with zero.

NA. Mixed model Poisson regression was used to assess the effect of *SLC6A4* and *HTR2A* on NA. Poisson regression was used because distributions of NA are right-skewed for all sampling periods and NA scores are discrete. Normally distributed varying intercepts for participants were included because of the repeated nature of the sampling. NA generally increased during the experiment and decreased at the final sampling period, NA5 (i.e., after the No Punishment game and at the end of the experimental session). Consistent with this observation, comparison of AIC weights indicates vastly more support for a model that specifies a parabolic relationship between sampling period and NA.

The effect of each locus on NA was explored separately. Three different specifications of the relationship between haplotype and phenotype (dominant, recessive, incomplete dominant) were considered. Interactions between haplotype and sampling period were considered as well. Model selection was conducted via AIC. For *SLC6A4*, the model specifying a main effect of *SH1*, with a recessive relationship between haplotype and phenotype, receives the greatest support. However, the 95% confidence intervals for the coefficient for *SH1* overlap considerably with zero.

For *HTR2A*, the model specifying a interaction between sampling period and *HH1*, with a recessive relationship between haplotype and phenotype, receives the greatest support. Inspection of the fixed effects coefficients for this model (Table S3) and predictions for NA from the fitted model (Fig. S2) reveals a growing trend for higher NA for *HH1* homozygotes following introduction of the No Punishment game (sampling period NA2).

Cortisol. Area under the curve with respect to ground (AUC_G) provides a measure of total cortisol secretion over a given time interval. AUC_G was calculated using equation 1 in ref. 48 for interval 1 (from T1 to T2, 15 min after the end of the No Punishment game) and for interval 2 (from T2 to T3, 15 min after the end of the Punishment game). The effect of variation at *HTR2A* and *SLC6A4* on AUC_G over intervals 1 and 2 was assessed with regression analyses. The natural logarithm of AUC_G was used in all analyses because distributions for this measurement were positively skewed. *HTR2A* and *SLC6A4* were analyzed separately, as were intervals 1 and 2. Cortisol response to a psychosocial stressor may be affected by sex, hormonal contraceptive use (49), and depression (50), and the effect of depression on cortisol reactivity may be particularly strong in the afternoon (50). Thus, a base model was constructed with the variables *Sex contraceptive* and *MDI*. *Sex contraceptive* consists of three categories: male, female, and female using hormonal contraceptive. *MDI* is the square root of the participant's score on the Major Depression Inventory. Haplotypes were added to this base model. All possible models allowing for interactions among these three variables (*Sex contraceptive*, *MDI*, and haplotype) were considered. This was done for three specifications of the relationship between haplotype and phenotype (dominant, recessive, incomplete dominant). AIC was used to select the best model for each specification of the relationship between haplotype and phenotype and then to select the best model among these three specifications.

For interval 1, the base model receives far more support than any of the models that include either *SLC6A4* or *HTR2A*. That is, there is no evidence that variation at *SLC6A4* or *HTR2A* affected total cortisol secretion during the No Punishment game. However, for interval 2, models that include *SLC6A4* or *HTR2A* outperform the base model. The best candidate models specify dominance of *SH1* and *HH1*. The best candidate model for *SLC6A4* does not specify interactions among any of the variables and indicates a negative effect of *SH1* on total cortisol secretion. Although AIC supports inclusion of *SLC6A4*, the SE of the coefficient for *SH1* is quite large relative to the point estimate and the 95% confidence interval overlaps with zero.

Comparison of the best candidate models for *SLC6A4* and *HTR2A* for total cortisol secretion over interval 2 indicates far more support for the model with *HTR2A*. The best candidate model for *HTR2A* includes a three-way interaction among *Sex contraceptive*, *MDI*, and *HH1* (Table S4). Total cortisol secretion during the Punishment game was greater for females with one or two copies of *HH1* but not for those with high scores on the Major Depression Inventory (Table S4).

ACKNOWLEDGMENTS. We thank G. Wright, G. Pepper, B. Verrelli, H. Abdolmaleky, and J. Quinn for helpful comments on previous versions of the manuscript. This work was funded by National Science Foundation Award 1003961 (to K.B.S.).

1. Boyd R, Richerson PJ (1992) Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethol Sociobiol* 195:171–195.
2. Sober E, Wilson DS (1999) *Do Unto Others: Evolution and Psychology of Unselfish Behavior* (Harvard Univ Press, Cambridge, MA).
3. Chudek M, Henrich J (2011) Culture-gene coevolution, norm-psychology and the emergence of human prosociality. *Trends Cogn Sci* 15(5):218–226.
4. Spitzer M, Fischbacher U, Herrnberger B, Grön G, Fehr E (2007) The neural signature of social norm compliance. *Neuron* 56(1):185–196.
5. Robinson OJ, Cools R, Sahakian BJ (2012) Tryptophan depletion disinhibits punishment but not reward prediction: Implications for resilience. *Psychopharmacology (Berl)* 219(2):599–605.
6. Crockett MJ, Clark L, Robbins TW (2009) Reconciling the role of serotonin in behavioral inhibition and aversion: Acute tryptophan depletion abolishes punishment-induced inhibition in humans. *J Neurosci* 29(38):11993–11999.
7. Crockett MJ, Clark L, Hauser MD, Robbins TW (2010) Serotonin selectively influences moral judgment and behavior through effects on harm aversion. *Proc Natl Acad Sci USA* 107(40):17433–17438.
8. Browning M, Holmes EA, Harmer CJ (2010) The modification of attentional bias to emotional information: A review of the techniques, mechanisms, and relevance to emotional disorders. *Cogn Affect Behav Neurosci* 10(1):8–20.
9. Crişan LG, et al. (2009) Genetic contributions of the serotonin transporter to social learning of fear and economic decision making. *Soc Cogn Affect Neurosci* 4(4):399–408.
10. Hariri AR, et al. (2002) Serotonin transporter genetic variation and the response of the human amygdala. *Science* 297(5580):400–403.
11. Pluess M, Belsky J, Way BM, Taylor SE (2010) 5-HTTLPR moderates effects of current life events on neuroticism: Differential susceptibility to environmental influences. *Prog Neuropsychopharmacol Biol Psychiatry* 34(6):1070–1074.
12. Fiocco AJ Joober R, Poirer J, Lupien S (2007) Polymorphism of the $5HT_{2A}$ receptor gene: Association with stress-related indices in healthy middle-aged adults. *Front Behav Neurosci* 1:3.
13. Haddley K, Bubb VJ, Breen G, Parades-Esquivel UM, Quinn JP (2012) Behavioural genetics of the serotonin transporter. *Curr Top Behav Neurosci* 12:503–535.
14. Serretti A, Drago A, De Ronchi D (2007) *HTR2A* gene variants and psychiatric disorders: A review of current literature and selection of SNPs for future studies. *Curr Med Chem* 14(19):2053–2069.
15. Way BM, Taylor SE (2010) The serotonin transporter promoter polymorphism is associated with cortisol response to psychosocial stress. *Biol Psychiatry* 67(5):487–492.
16. Claw KG, Tito RY, Stone AC, Verrelli BC (2010) Haplotype structure and divergence at human and chimpanzee serotonin transporter and receptor genes: Implications for behavioral disorder association analyses. *Mol Biol Evol* 27(7):1518–1529.
17. Fehr E, Gächter S (2000) Cooperation and punishment in Public Goods experiments. *Am Econ Rev* 90:980–994.
18. Herrmann B, Thöni C, Gächter S (2008) Antisocial punishment across societies. *Science* 319(5868):1362–1367.

19. Ashley R, Ball SB, Eckel CC (2010) Motives for giving: A reanalysis of two classic public goods experiments. *South Econ J* 77:15–26.
20. Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Automat Contr* 19:716–723.
21. Boulougouris V, Glennon JC, Robbins TW (2008) Dissociable effects of selective $5HT_{2A}$ and $5\text{-}HT_{2C}$ receptor antagonists on serial spatial reversal learning in rats. *Neuropsychopharmacology* 33(8):2007–2019.
22. Gächter S, Thöni C (2005) Social learning and voluntary cooperation among like-minded people. *J Eur Econ Assoc* 3:303–314.
23. Rustagi D, Engel S, Kosfeld M (2010) Conditional cooperation and costly monitoring explain success in forest commons management. *Science* 330(6006):961–965.
24. Dinan TG (1996) Serotonin and the regulation of hypothalamic-pituitary-adrenal axis function. *Life Sci* 58(20):1683–1694.
25. Chaouloff F (2000) Serotonin, stress and corticoids. *J Psychopharmacol* 14(2):139–151.
26. Henrich J, et al. (2006) Costly punishment across human societies. *Science* 312(5781): 1767–1770.
27. Polesskaya OO, Aston C, Sokolov BP (2006) Allele C-specific methylation of the $5HT_{2A}$ receptor gene: Evidence for correlation with its expression and expression of DNA methylase *DNMT1*. *J Neurosci Res* 83(3):362–373.
28. Falkenberg VR, Gurbaxani BM, Unger ER, Rajeevan MS (2011) Functional genomics of serotonin receptor 2A (*HTR2A*): Interaction of polymorphism, methylation, expression and disease association. *Neuromolecular Med* 13(1):66–76.
29. Ali FR, et al. (2010) Combinatorial interaction between two human serotonin transporter gene variable number tandem repeats and their regulation by CTCF. *J Neurochem* 112(1):296–306.
30. Philibert R, et al. (2007) Serotonin transporter mRNA levels are associated with the methylation of an upstream CpG island. *Am J Med Genet B Neuropsychiatr Genet* 144B(1):101–105.
31. Keltikangas-Järvinen L, Salo J (2009) Dopamine and serotonin systems modify environmental effects on human behavior: A review. *Scand J Psychol* 50(6):574–582.
32. Belsky J, et al. (2009) Vulnerability genes or plasticity genes? *Mol Psychiatry* 14(8): 746–754.
33. Laland KN, Odling-Smee J, Myles S (2010) How culture shaped the human genome: Bringing genetics and the human sciences together. *Nat Rev Genet* 11(2):137–148.
34. Richerson PJ, Boyd R, Henrich J (2010) Colloquium paper: Gene-culture coevolution in the age of genomics. *Proc Natl Acad Sci USA* 107(Suppl 2):8985–8992.
35. Chiao JY, Blizinsky KD (2010) Culture-gene coevolution of individualism-collectivism and the serotonin transporter gene. *Proc Biol Sci* 277(1681):529–537.
36. Bell AM, Robinson GE (2011) Genomics. Behavior and the dynamic genome. *Science* 332(6034):1161–1162.
37. Bech P, Rasmussen NA, Olsen LR, Noerholm V, Abildgaard W (2001) The sensitivity and specificity of the Major Depression Inventory, using the Present State Examination as the index of diagnostic validity. *J Affect Disord* 66(2-3):159–164.
38. Thompson ER (2007) Development and validation of an internationally reliable short-form of the Positive and Negative Affect Schedule (PANAS). *J Cross Cult Psychol* 38:227–242.
39. Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68(4):978–989.
40. Stephens M, Scheet P (2005) Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am J Hum Genet* 76(3):449–462.
41. Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet* 73(5):1162–1169.
42. R Development Core Team (2011) R: A language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna), Version 2.14.1.
43. McElreath R (2012) rethinking: Statistical Rethinking Book Package, R Package Version 1.04 (Univ of California, Davis, CA).
44. Bates D, Maechler M, Bolker B (2011) lme4: Linear Mixed-Effects Models Using S4 Classes, R Package Version 0.999375-42. Available at http://CRAN.R-project.org/package=lme4.
45. McElreath R (2012) glmer2stan: Rstan models defined by glmer formulas, R package version 0.96 (Univ of California, Davis, CA).
46. Wickham H (2009) *ggplot2: Elegant Graphics for Data Analysis* (Springer, New York).
47. Stan Development Team (2012) Stan: A C++ Library for Probability and Sampling, Version 1.0. Available at http://mc-stan.org/.
48. Pruessner JC, Kirschbaum C, Meinlschmid G, Hellhammer DH (2003) Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. *Psychoneuroendocrinology* 28(7):916–931.
49. Kudielka BM, Hellhammer DH, Wüst S (2009) Why do we respond so differently? Reviewing determinants of human salivary cortisol responses to challenge. *Psychoneuroendocrinology* 34(1):2–18.
50. Burke HM, Davis MC, Otte C, Mohr DC (2005) Depression and cortisol responses to psychological stress: A meta-analysis. *Psychoneuroendocrinology* 30(9):846–856.