

6-15-2015

Genomic Signatures of Cooperation and Conflict in the Social Amoeba

Elizabeth A. Ostrowski

Yufeng Shen

Xiangjun Tian

Richard Sucgang

Huaiyang Jiang

See next page for additional authors

Follow this and additional works at: http://openscholarship.wustl.edu/bio_facpubs

 Part of the [Behavior and Ethology Commons](#), [Biology Commons](#), [Environmental Microbiology and Microbial Ecology Commons](#), and the [Population Biology Commons](#)

Recommended Citation

Ostrowski, Elizabeth A.; Shen, Yufeng; Tian, Xiangjun; Sucgang, Richard; Jiang, Huaiyang; Qu, Jiaxin; Katoh-Kurasawa, Mariko; Brock, Debra A.; Dinh, Christopher; Lara-Garduno, Fremiet; Lee, Sandra L.; Kovar, Christie L.; Dinh, Huyen H.; Korchina, Viktoriya; Jackson, LaRonda; Patil, Shobha; Han, Yi; Chaboub, Lesley; Shaulsky, Gad; Muzny, Donna M.; Worley, Kim C.; Gibbs, Richard A.; Richards, Stephen; Kuspa, Adam; Strassmann, Joan E.; and Queller, David C., "Genomic Signatures of Cooperation and Conflict in the Social Amoeba" (2015). *Biology Faculty Publications & Presentations*. Paper 68.
http://openscholarship.wustl.edu/bio_facpubs/68

This Article is brought to you for free and open access by the Biology at Washington University Open Scholarship. It has been accepted for inclusion in Biology Faculty Publications & Presentations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

Authors

Elizabeth A. Ostrowski, Yufeng Shen, Xiangjun Tian, Richard Sucgang, Huaiyang Jiang, Jiaxin Qu, Mariko Katoh-Kurasawa, Debra A. Brock, Christopher Dinh, Fremiet Lara-Garduno, Sandra L. Lee, Christie L. Kovar, Huyen H. Dinh, Viktoriya Korchina, LaRonda Jackson, Shobha Patil, Yi Han, Lesley Chaboub, Gad Shaulsky, Donna M. Muzny, Kim C. Worley, Richard A. Gibbs, Stephen Richards, Adam Kuspa, Joan E. Strassmann, and David C. Queller

1 **Genomic Signatures of Cooperation and Conflict in the Social Amoeba**

2

3 Elizabeth A. Ostrowski^{1*}, Yufeng Shen², Xiangjun Tian³, Richard Sucgang⁴, Huaiyang Jiang⁵,
4 Jiaxin Qu⁵, Mariko Katoh-Kurasawa⁶, Debra A. Brock⁷, Christopher Dinh⁴, Fremiet Lara-
5 Garduno⁵, Sandra L. Lee⁵, Christie L. Kovar⁵, Huyen H. Dinh⁵, Viktoriya Korchina⁵, LaRonda
6 Jackson⁵, Shobha Patil⁵, Yi Han⁵, Lesley Chaboub⁵, Gad Shaulsky⁶, Donna M. Muzny^{5,6}, Kim C.
7 Worley^{5,6}, Richard A. Gibbs^{5,6}, Stephen Richards⁵, Adam Kuspa⁴, Joan E. Strassmann⁷, and
8 David C. Queller⁷.

9

10 *Author for correspondence

11

12 ¹Department of Biology and Biochemistry, University of Houston, Houston, TX 77204, USA

13 ²Departments of Systems Biology and Biomedical Informatics, Columbia University, New York,
14 NY 10027, USA

15 ³Alkek Center for Metagenomics and Microbiome Research, Department of Molecular Virology
16 and Microbiology, Baylor College of Medicine, Houston, TX 77030, USA

17 ⁴Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of
18 Medicine, Houston, TX 77030, USA

19 ⁵Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX 77030 USA

20 ⁶Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030
21 USA

22 ⁷Department of Biology, Washington University in St. Louis, St. Louis, MO 63130, USA

23

24 *Contact Information: eaostrowski@uh.edu

25

1 **Running Title**

2 Genomic Signatures of Cooperation and Conflict

3

4 **Summary**

5 Cooperative systems are susceptible to invasion by selfish individuals that profit from receiving
6 the social benefits but fail to contribute. These so-called “cheaters” can have a fitness
7 advantage in the laboratory, but it is unclear whether cheating provides an important selective
8 advantage in nature. We used a population genomic approach to examine the history of genes
9 involved in cheating behaviors in the social amoeba *Dictyostelium discoideum*, testing whether
10 these genes experience rapid evolutionary change as a result of conflict over spore-stalk fate.
11 Candidate genes and surrounding regions showed elevated polymorphism, unusual patterns of
12 linkage disequilibrium, and lower levels of population differentiation, but they did not show
13 greater between-species divergence. The signatures were most consistent with frequency-
14 dependent selection acting to maintain multiple alleles, suggesting that conflict may lead to
15 stalemate rather than an escalating arms race. Our results reveal the evolutionary dynamics of
16 cooperation and cheating and underscore how sequence-based approaches can be used to
17 elucidate the history of conflicts that are difficult to observe directly.

18

19 **Highlights**

- 20 • Molecular evolution analyses reveal the history of social conflict
- 21 • Genes that mediate social conflict show signatures of frequency-dependent selection
- 22 • Balanced polymorphisms suggest that cheating may be stable and endemic

23

24

25 **Results**

1 The social amoeba *Dictyostelium discoideum* is a model system for cooperation and conflict [1,
2 2]. Upon starvation, up to hundreds of thousands of amoebae converge, forming a multicellular
3 slug and eventually a fruiting body. Some cells form the stalk of the fruiting body and die, while
4 others rise to the top, form resistant spores, and disperse. Stalk formation is altruistic, because
5 death of the stalk cells enhances the survival and dispersal of the spores. However, because
6 fruiting bodies can contain multiple, genetically different clones, selection can favor cheaters –
7 individuals that avoid forming the stalk themselves, yet benefit from its production by others [2].
8
9 Consistent with the prediction of cheating, cheaters can be readily found in natural populations
10 of *D. discoideum* [1-3]. However, whether individuals that cheat are evolutionarily successful is
11 unclear, and several hypotheses have been proposed. One hypothesis is that cheating selects
12 for resistance, and resistance in turn selects for greater cheating. Social conflict could thus drive
13 an escalating arms race of adaptations and counter-adaptations, reminiscent of the arms races
14 between hosts and pathogens or predators and prey [4-7]. An alternative possibility is that
15 cheaters have a selective advantage only when rare. For example, as cheaters increase in
16 frequency in a population, they potentially displace the very victims on which they depend, or
17 face other trade-offs [8]. This negative frequency-dependence predicts that cheaters and
18 cooperators can be maintained as a balanced polymorphism, effectively leading to a stalemate
19 (Fig. 1B) [9, 10]. Finally, some have suggested that there is no selective advantage to cheating
20 [11]. Cheating might be selected against if relatedness among the strains in a fruiting body is
21 high, such that cheaters primarily cheat their own relatives [12]. In this case, cheating might
22 persist in populations as a ‘cheating load’, analogous to a genetic load for deleterious mutations
23 (Fig. 1C). Alternatively, cheating might also not be favored if the multicellular stage occurs only
24 rarely in nature [13], such that there is little selection for or against these phenotypes.

25

1 Crucially, these different hypotheses about the long-term success of social cheating make
2 unique, testable predictions about variation in the genes that mediate these conflicts ([14]; Table
3 1). To distinguish among these different possibilities, we took advantage of a previous screen
4 that identified over >150 loci in *D. discoideum* that impact cheating behaviors [5]. We used
5 whole genome sequencing and molecular evolution to ask whether genes that mediate cheating
6 behaviors show distinctive signatures of molecular evolution that differ from the rest of the
7 genome, distinguishing among the different hypotheses described in Table 1.

8

9 **Candidate Genes Show Elevated Polymorphism**

10 The different evolutionary scenarios for cheating alleles make unique predictions about the
11 levels of polymorphism versus divergence (Table 1). For example, an escalating arms race
12 driven by repeated selective sweeps of cheating alleles should reduce variation within species
13 while elevating the sequence divergence between species, whereas the stalemate model
14 makes the opposite prediction. To test these possibilities, we first examined levels of
15 polymorphism in regions surrounding candidate genes, comparing these values to a null
16 hypothesis based on other regions of the genome. We observed higher polymorphism in
17 candidate genes as a group compared to randomly chosen regions, which was significant for
18 both mean and median levels at 20 kb (Fig. 2). Compared to other genes, sequence variation
19 was also disproportionately non-synonymous (higher p_N/p_S ; Table S1). Higher levels of
20 polymorphism might occur if genes important for cheating behaviors show limited expression
21 (e.g., if they are expressed in an infrequent portion of the life cycle [13]), but analysis of the
22 published transcriptome of the lab strain [15] indicated no difference in their timing or levels of
23 expression compared to other genes in the genome (Table S2).

24

25 **No Evidence of Elevated Sequence Divergence between Species**

1 The arms race hypothesis also predicts elevated rates of divergence between species, so we
2 compared *D. discoideum* to its sister species *D. citrinum* at all identified orthologs. These
3 analyses revealed lower rates of non-synonymous to synonymous substitution (d_N/d_S) for
4 candidate genes compared to other genes in the genome (Table S1), which was inconsistent
5 with the predictions of an arms race (Table 1). We reached a similar conclusion using
6 McDonald-Kreitman tests [16], which compare p_N/p_S to d_N/d_S for each gene individually. Given
7 very low levels of sequence polymorphism within *D. discoideum*, there was little resolution to
8 detect significant deviations in this ratio for each locus individually. Nevertheless, two candidate
9 genes (DDB_G0285541 and *chtC*) had strongly significant McDonald-Kreitman tests that were
10 also in the extreme tail of the genome-wide distribution. These genes showed elevations in d_N/d_S
11 relative to p_N/p_S , indicative of directional selection driving sequence divergence. In the remainder
12 of candidate genes (that is, removing these two genes), the ratio was strongly significant and
13 opposite in direction, indicating an excess of non-synonymous polymorphism to non-
14 synonymous divergence compared to other genes (Fig. 3). Taken together, genes mediating
15 social cheating did not show the elevated rates of amino-acid substitution predicted under an
16 escalating arms race or relaxed selection scenario.

17

18 **Additional Signatures of Selection Support Stalemates**

19 At first glance, elevated non-synonymous polymorphism, combined with low amino acid
20 divergence, is potentially consistent with the stalemate model of frequency-dependent selection,
21 where novel alleles can invade and establish in populations, but ultimately fail to take over.
22 Several additional tests support this interpretation. First, scaled to gene length, we observed
23 significantly fewer haplotypes than expected and high levels of intragenic linkage disequilibrium
24 (lower \bullet ; Fig. S1), indicating divergent alleles, a signature of balancing selection. In addition, two

1 metrics of balancing selection, Wall's B and Wall's Q [17], were significantly elevated in
2 sequence windows surrounding candidate loci (Table S3).

3
4 A common test for balancing selection is to examine the distribution of allele frequencies –
5 whereas positive or purifying selection produce a strongly skewed distribution, balancing
6 selection can maintain multiple alleles at intermediate frequencies. Surprisingly, given our
7 results above supporting balancing selection, candidate loci showed greater skew, indicated by
8 more negative values for two metrics of the site frequency spectrum (Table S4). Candidate
9 genes as a group also showed a significant excess of high frequency derived alleles (Fay and
10 Wu's H : -0.002; $P=0.03$), which can indicate incomplete selective sweeps. The significant,
11 negative Fay and Wu's H test for candidate genes suggests that variants rise to high frequency
12 quickly and that these genes experience stronger-than-expected selection for sequence
13 changes.

14
15 Finally, we calculated the index of population structure (F_{ST}) at each segregating site in the
16 genome, comparing SNPs in close proximity to candidate genes against the rest of the genome.
17 Unusually high F_{ST} can indicate that different alleles predominate in different geographic
18 locations (in this case, Texas and Virginia) and can be caused by geographically restricted
19 selective sweeps. On the other hand, unusually low F_{ST} values indicate alleles attain similar
20 frequencies across geographically distinct populations, with each subpopulation maintaining
21 multiple divergent alleles—a signature of negative frequency-dependent selection [18]. These
22 results revealed lower-than-expected F_{ST} at candidate compared to non-candidate loci (Fig. S2).
23 The elevation in polymorphism in these genes, combined with significant reductions in
24 population structure, argues against divergent alleles caused by local sweeps and suggests a

1 role for negative frequency-dependent selection, with selection maintaining the same variants
2 across subpopulations. •

3

4 **Evidence for Multiple Modes of Selection**

5 By examining the molecular evolution patterns of candidate genes as a group, we could
6 determine whether genes that mediate social conflict have general, recognizable patterns
7 indicating unique forms of selection. Analyzing these genes as a group also provided greater
8 sensitivity, which was helpful given the low levels of polymorphism in this species. However,
9 these results reflect only average differences between candidate and non-candidate genes –
10 and in fact, a diversity of dynamics is possible. Moreover, for many metrics, candidate genes
11 might be extreme in opposite directions, such that we may fail to observe a strong signal of
12 selection because these effects average out. We addressed this possibility in two ways. First,
13 for each evolutionary metric, we asked whether the variance was higher for candidate genes
14 compared to groups where genes are chosen randomly, indicating extremes in opposite
15 directions. However, we did not observe elevated variance for any metric (Table S5). We also
16 asked whether there was overrepresentation of candidate genes in both tails of the genome-
17 wide distribution for each metric, but we observed no such cases (Table S6). While there was
18 no overrepresentation in either tail of the genome-wide distribution for the McDonald-Kreitman
19 test, two genes (DDB_G0285541 and *chtC*) showed extreme signatures of positive (directional)
20 selection. The remaining genes showed the opposite pattern, an average excess of within-
21 species non-synonymous polymorphism compared to non-synonymous divergence. Thus, while
22 the average signature in conflict-related genes was elevated polymorphism and other patterns
23 suggesting the selective maintenance of multiple alleles, other signatures were evident as well,
24 including signatures of strong directional selection that could reflect escalating arms races at
25 these particular loci.

1

2 **Discussion**

3 In *D. discoideum*, a model system for the study of social conflict, it has long been noted that
4 cheating behaviors are present among natural isolates [2, 3], but there is little understanding of
5 why selfish behaviors arise, whether they persist, and whether the prevalence of cheaters in
6 natural populations signifies long-term evolutionary success of this social strategy. We have
7 indications of three main signatures of balancing selection on these genes as a class compared
8 to other genes. First, they show higher levels of polymorphism, as expected when at least one
9 SNP is under balancing selection and increases diversity at linked neutral loci. However, they
10 did not show the concomitant elevation in amino acid substitution expected if polymorphism
11 simply reflected relaxed selection. Second, they showed lower F_{ST} values, which can occur if
12 selection is maintaining the same balanced polymorphisms across geographically different
13 subpopulations. Third, another indicator of the deeper coalescence times characteristic of
14 balancing selection is the significantly fewer haplotypes and higher linkage disequilibrium values,
15 including significantly elevated Wall's B and Wall's Q .

16

17 The observation of elevated levels of polymorphism surrounding genes implicated in social
18 cheating, combined with other signatures of recent selection, argue that *D. discoideum*
19 experiences ongoing selection at these loci and is consistent with frequency-dependent
20 selection allowing multiple types (or alleles) to coexist. Notably, we failed to observe the
21 molecular signatures of a simple arms race – these dynamics are expected to reduce genetic
22 diversity and drive long-term sequence divergence between populations and/or species. Rather,
23 our results are reminiscent of ‘trench warfare’, an alternative arms race scenario where
24 alternative alleles do not rapidly displace one another, resulting in a prolonged stalemate [19].

1 Negative-frequency dependence is also a hallmark of Red Queen dynamics, a form of
2 coevolutionary arms race where alleles continually cycle but rarely fix [20].
3
4 The finding of balancing selection is also consistent with evolutionary theory about the role of
5 frequency-dependence in social interactions [21]. Many social behaviors are inherently
6 frequency-dependent, where the fitness of a given strategy (e.g., cheat or cooperate) is
7 dependent on whether an individual's social partners employ the same strategy or not.
8 Experimental studies of bacterial mutants that exhibit cheating behaviors suggest frequency-
9 dependence might be common [22-26]. In *Dictyostelium*, frequency-dependent fitness was
10 shown for the *fbxA-* strain, a mutant that allocates fewer cells to the stalk but produces
11 disproportionately many spores when co-developed with another strain, which it cheats [12].
12
13 Like social conflict, conflicts between the sexes over optimal levels of mating and between
14 parents and offspring over optimal provisioning are also hypothesized to result in antagonistic
15 coevolution [20, 27-29]. While some studies have shown directional selection on genes
16 underlying these other forms of intraspecific conflict, many others have found signatures of
17 balancing selection, diversifying selection, or a combination of both [30-33]. Thus taken together,
18 our results not only indicate stalemate as a possible outcome of social conflict, but add to a
19 growing body of evidence that stalemates may be a common outcome in conflict-driven systems
20 more generally. While identification of these polymorphisms should open the door to
21 investigation into the functional consequences of this variation for cheating and resistance
22 behaviors, the population genomic approach used here provides insight into the long-term
23 consequences of social conflict and highlights the possibility of an ongoing, dynamic interaction
24 at these loci.

25

1 **Experimental Procedures**

2 **Strains.** We re-sequenced the genomes of 20 natural isolates of *D. discoideum*, primarily from
3 two locations: Houston, TX (six strains) and Mountain Lake, VA (nine strains), as well as 5
4 additional strains from different geographic locations (two sites in Texas, and one site in each of
5 Massachusetts, Kentucky, and Illinois) using 454 or Illumina. Sequencing reads were aligned to
6 the Ax4 reference genome (Assembly/GFF3 file generated June 9 2010, available at
7 dictybase.org) using MAQ for Illumina sequencing reads and ATLAS-SNP for 454 data. Detailed
8 mapping and SNP calling procedures are available in the Supplemental Information.

9

10 **Molecular Evolution Analyses.** Nucleotide diversity, Tajima's *D*, Fu and Li's *D*^{*}, Hudson's *C*
11 (recombination, or rho), haplotype diversity, Fay and Wu's *H*, and haplotype number were
12 determined for all genes in the genome and in sequence windows using "compute" (available at
13 molpopgen.org). Levels of non-synonymous (p_n) and synonymous (p_s) diversity were calculated
14 using the program "gestimator", and the McDonald-Kreitman tests were obtained using "MKtest"
15 (both available at molpopgen.org). F_{ST} was calculated for all segregating sites using scripts
16 written in Ruby and Python. Resampling analyses were performed using R. More details are
17 available in the Supplemental Information.

18

19 **Author Contributions**

20 Cell culture and DNA extraction: EAO, CD, MKK, DB; Conceived and designed the experiments:
21 EAO, JES, DCQ, SR, KCW, RAG, RS; Library preparation and sequencing: FL-G, SLL, CK, HD,
22 VK, LJ, SP, YH, LC, DMM, SR, RAG, KW; Analyzed the data: EAO, YS, XT, RS, SR; Assembly
23 and annotation of *D. citrinum*: JQ, HJ, KW; Analysis of *D. citrinum* and *D. discoideum*: EAO, XT;
24 Wrote the paper: EAO, JES, AK, DCQ.

25

1 **Acknowledgements**

2 We thank Rasmus Nielsen for discussions of the data and comments on an earlier version of
3 this manuscript.

1 **References**

- 2 1. Strassmann, J. E., and Queller, D. C. (2011). Evolution of cooperation and control of
3 cheating in a social microbe. *Proc. Natl. Acad. Sci. U.S.A.* *108 Suppl 2*, 10855–10862.
- 4 2. Strassmann, J. E., Zhu, Y., and Queller, D. C. (2000). Altruism and social cheating in the
5 social amoeba *Dictyostelium discoideum*. *Nature* *408*, 965–967.
- 6 3. Buttery, N. J., Rozen, D. E., Wolf, J. B., and Thompson, C. R. L. (2009). Quantification of
7 social behavior in *D. discoideum* reveals complex fixed and facultative strategies. *Curr.*
8 *Biol.* *19*, 1373–1377.
- 9 4. Khare, A., Santorelli, L. A., Strassmann, J. E., Queller, D. C., Kuspa, A., and Shaulsky, G.
10 (2010). Cheater-resistance is not futile. *Nature* *461*, 980–982.
- 11 5. Santorelli, L. A., Thompson, C. R. L., Villegas, E., Svez, J., Dinh, C., Parikh, A., Sucgang,
12 R., Kuspa, A., Strassmann, J. E., Queller, D. C., et al. (2008). Facultative cheater
13 mutants reveal the genetic complexity of cooperation in social amoebae. *Nature* *451*,
14 1107–1110.
- 15 6. Hollis, B. (2012). Rapid antagonistic coevolution between strains of the social amoeba
16 *Dictyostelium discoideum*. *Proc. Biol. Sci.* *279*, 3565–3571.
- 17 7. Zhang, Q.-G., Buckling, A., Ellis, R. J., and Godfray, H. C. J. (2009). Coevolution
18 between cooperators and cheats in a microbial system. *Evolution* *63*, 2248–2256.
- 19 8. Sathe, S., Khetan, N., and Nanjundiah, V. (2014). Interspecies and intraspecies
20 interactions in social amoebae. *J. Evol. Biol.* *27*, 349–362.
- 21 9. Maynard Smith, J. (1982). *Evolution and the Theory of Games* (Cambridge: Cambridge
22 University Press).
- 23 10. Doebeli, M., Hauert, C., and Killingback, T. (2004). The evolutionary origin of cooperators
24 and defectors. *Science* *306*, 859–862.
- 25 11. Van Dyken, J. D., Linksvayer, T. A., and Wade, M. J. (2011). Kin selection–mutation
26 balance: A model for the origin, maintenance, and consequences of social cheating. *Am.*
27 *Nat.* *177*, 288–300.
- 28 12. Gilbert, O. M., Foster, K. R., Mehdiabadi, N. J., Strassmann, J. E., and Queller, D. C.
29 (2007). High relatedness maintains multicellular cooperation in a social amoeba by
30 controlling cheater mutants. *Proc. Natl. Acad. Sci. U.S.A.* *104*, 8913–8917.
- 31 13. Linksvayer, T. A., and Wade, M. J. (2009). Genes with social effects are expected to
32 harbor more sequence variation within and between species. *Evolution* *63*, 1685–1696.
- 33 14. Van Dyken, J. D., and Wade, M. J. (2012). Detecting the molecular signature of social
34 conflict: theory and a test with bacterial quorum sensing genes. *Am. Nat.* *179*, 436–450.
- 35 15. Parikh, A., Miranda, E. R., Katoh-Kurasawa, M., Fuller, D., Rot, G., Zagar, L., Curk, T.,
36 Sucgang, R., Chen, R., Zupan, B., et al. (2010). Conserved developmental

- 1 transcriptomes in evolutionarily divergent species. *Genome Biol.* *11*, R35.
- 2 16. McDonald, J. H., and Kreitman, M. (1991). Adaptive protein evolution at the *Adh* locus in
3 *Drosophila*. *351*, 652–654.
- 4 17. Wall, J. D. (1999). Recombination and the power of statistical tests of neutrality. *74*, 65–
5 79.
- 6 18. Glémin, S. (2005). Balancing selection in the wild: testing population genetics theory of
7 self-incompatibility in the rare species *Brassica insularis*. *Genetics* *171*, 279–289.
- 8 19. Stahl, E. A., Dwyer, G., Mauricio, R., Kreitman, M., and Bergelson, J. (1999). Dynamics
9 of disease resistance polymorphism at the *Rpm1* locus of *Arabidopsis*. *Nature* *400*, 667–
10 671.
- 11 20. Brockhurst, M. A., Chapman, T., King, K. C., Mank, J. E., Paterson, S., and Hurst, G. D.
12 D. (2014). Running with the Red Queen: the role of biotic conflicts in evolution. *Proc. Biol.*
13 *Sci.* *281*, 20141382.
- 14 21. Queller, D. C. (2011). Expanded social fitness and Hamilton's rule for kin, kith, and kind.
15 *Proc. Natl. Acad. Sci. U.S.A.* *108*, 10792–10799.
- 16 22. Diggle, S. P., Griffin, A. S., Campbell, G. S., and West, S. A. (2007). Cooperation and
17 conflict in quorum-sensing bacterial populations. *Nature* *450*, 411–414.
- 18 23. Ross-Gillespie, A., Gardner, A., West, S. A., and Griffin, A. S. (2007). Frequency
19 dependence and cooperation: theory and a test with bacteria. *Am. Nat.* *170*, 331–342.
- 20 24. Dugatkin, L. A., Perlin, M., Lucas, J. S., and Atlas, R. (2005). Group-beneficial traits,
21 frequency-dependent selection and genotypic diversity: an antibiotic resistance paradigm.
22 *Proc. Biol. Sci.* *272*, 79–83.
- 23 25. Ellis, R. J., Lilley, A. K., Lacey, S. J., Murrell, D., and Godfray, H. C. J. (2007).
24 Frequency-dependent advantages of plasmid carriage by *Pseudomonas* in homogeneous
25 and spatially structured environments. *ISME J* *1*, 92–95.
- 26 26. MacLean, R. C., and Gudelj, I. (2006). Resource competition and social conflict in
27 experimental populations of yeast. *Nature* *441*, 498–501.
- 28 27. Arnqvist, G., and Rowe, L. (2005). *Sexual Conflict* (Princeton: Princeton University Press).
- 29 28. Rice, W. R., and Holland, B. (1997). The enemies within: intergenomic conflict, interlocus
30 contest evolution (ICE), and the intraspecific Red Queen. *Behav. Ecol. Sociobiol.* *41*, 1–
31 10.
- 32 29. Trivers, R. L. (1974). Parent-offspring conflict. *Am. Zool.* *14*, 249–264.
- 33 30. Hamm, D., Mautz, B. S., Wolfner, M. F., Aquadro, C. F., and Swanson, W. J. (2007).
34 Evidence of amino acid diversity-enhancing selection within humans and among
35 primates at the candidate sperm-receptor gene *PKDREJ*. *Am. J. Hum. Gen.* *81*, 44–52.

- 1 31. Kelleher, E. S., Clark, N. L., and Markow, T. A. (2011). Diversity-enhancing selection acts
2 on a female reproductive protease family in four subspecies of *Drosophila mojavensis*.
3 *Genetics* 187, 865–876.
- 4 32. Kawabe, A., Fujimoto, R., and Charlesworth, D. (2007). High diversity due to balancing
5 selection in the promoter Region of the *Medea* Gene in *Arabidopsis lyrata*. *Curr. Biol.* 17,
6 1885–1889.
- 7 33. Miyake, T., Takebayashi, N., and Wolf, D. E. (2009). Possible diversifying selection in the
8 imprinted gene, *MEDEA*, in *Arabidopsis*. *Mol. Biol. Evol.* 26, 843–857.
- 9
- 10

1 **Figure Legends**

2

3 **Figure 1. Example scenarios for the evolutionary dynamics of cheating behaviors.**

4 Shaded areas are proportional to the frequencies of different alleles (colors) in a population. (A)

5 Escalating arms races, where epidemics of cheating and/or resistance sweep through

6 populations, (B) Stalemates, where cheaters invade but neither fix nor become lost from the

7 population, resulting in endemic cheating, (C) Mutation-selection balance, where new mutations

8 that produce cheating behaviors are continually introduced into a population, but selection

9 removes them.

10

11 **Figure 2. Comparison of polymorphism in regions surrounding candidate loci compared**

12 **to the rest of the genome.** Dotted lines show the mean or median nucleotide diversity in

13 sequence windows of 10- or 20-kb compared to the null distribution based on 10,000 data sets

14 of the same size where sequence windows were chosen randomly. Asterisks indicate

15 statistically significant results. (A) Mean nucleotide diversity in 10-kb sequence windows, (B)

16 mean nucleotide diversity in 20-kb sequence windows, (C) median nucleotide diversity in 10-kb

17 sequence windows, and (D) median nucleotide diversity in 20-kb sequence windows.

18

19 **Figure 3. McDonald-Kreitman tests.** Histograms show the mean or median ratio of d_N/d_S to

20 p_N/p_S for 10,000 randomly generated gene sets, and the dotted line shows the observed value of

21 this ratio for candidate genes. Asterisks indicate statistically significant results. (A) Mean ratio,

22 (B) median ratio, (C and D) mean and median ratios after removing two candidate genes with

23 extreme values. See text for details.

24

25

1 **Tables**

2 **Table 1. Predicted sequence patterns for cooperation and cheating genes under different**
 3 **evolutionary scenarios.**

Scenario	Description	Within-species polymorphism	Between-species divergence	Additional Signatures
Escalating Arms Race (Directional Selection)	Repeated selective sweeps of cheating alleles through populations remove variation within populations and drive rapid divergence between species	↓	↑	Elevated population structure (higher F_{ST}), excess of high frequency derived alleles (negative Fay and Wu's H)
Stalemate (Balancing Selection)	Negative frequency-dependence maintains both cheaters and cooperators within populations	↑	↓	Reduced population structure (lower F_{ST}), excess of intermediate frequency alleles (positive Tajima's D), elevated linkage disequilibrium (lower r^2), haplotype structure (higher Wall's B and Wall's Q)
Relaxed Selection	Cheating behaviors do not experience strong selection, possibly	↑	↑	Allele frequency skew closer to zero (Tajima's $D=0$)

	because the multicellular (social) stage is rare in nature			
Purifying Selection	Cheaters are selected against, for example if relatedness is high	↓	↓	Excess of low frequency alleles (negative Tajima's <i>D</i>)

1
2