



## The Spread of Behavior in an Online Social Network Experiment

Damon Centola

*Science* **329**, 1194 (2010);

DOI: 10.1126/science.1185231

*This copy is for your personal, non-commercial use only.*

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

**The following resources related to this article are available online at [www.sciencemag.org](http://www.sciencemag.org) (this information is current as of October 17, 2012 ):**

**Updated information and services**, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/329/5996/1194.full.html>

**Supporting Online Material** can be found at:

<http://www.sciencemag.org/content/suppl/2010/08/31/329.5996.1194.DC1.html>

A list of selected additional articles on the Science Web sites **related to this article** can be found at:

<http://www.sciencemag.org/content/329/5996/1194.full.html#related>

This article **cites 20 articles**, 5 of which can be accessed free:

<http://www.sciencemag.org/content/329/5996/1194.full.html#ref-list-1>

This article has been **cited by** 12 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/content/329/5996/1194.full.html#related-urls>

This article appears in the following **subject collections**:

Sociology

<http://www.sciencemag.org/cgi/collection/sociology>

G. Hunt, A. Miller, T. Olszewski, and P. Wagner for their suggestions; M. Kosnik and A. Miller for reviews; and M. Foote for verifying that my subsampling algorithms were programmed correctly. Numerous contributors to the Paleobiology Database made this study possible, and I am particularly grateful to M. Clapham, A. Hendy, and W. Kiessling for recent contributions. Research described

here was funded by donations from anonymous private individuals having no connection to it. This is Paleobiology Database publication 117.

#### Supporting Online Material

www.sciencemag.org/cgi/content/full/329/5996/1191/DC1  
Materials and Methods

Figs. S1 to S9  
Tables S1 and S2  
References

22 March 2010; accepted 30 June 2010  
10.1126/science.1189910

# The Spread of Behavior in an Online Social Network Experiment

Damon Centola

How do social networks affect the spread of behavior? A popular hypothesis states that networks with many clustered ties and a high degree of separation will be less effective for behavioral diffusion than networks in which locally redundant ties are rewired to provide shortcuts across the social space. A competing hypothesis argues that when behaviors require social reinforcement, a network with more clustering may be more advantageous, even if the network as a whole has a larger diameter. I investigated the effects of network structure on diffusion by studying the spread of health behavior through artificially structured online communities. Individual adoption was much more likely when participants received social reinforcement from multiple neighbors in the social network. The behavior spread farther and faster across clustered-lattice networks than across corresponding random networks.

Many behaviors spread through social contact (1–3). As a result, the network structure of who is connected to whom can critically affect the extent to which a behavior diffuses across a population (2–8). There are two competing hypotheses about how network structure affects diffusion. The “strength of weak ties” hypothesis predicts that networks with many “long ties” (e.g., “small-world” topologies) will spread a social behavior farther and more quickly than a network in which ties are highly clustered (4–6). This hypothesis treats the spread of behavior as a simple contagion, such as disease or information: A single contact with an “infected” individual is usually sufficient to transmit the behavior (2). The power of long ties is that they reduce the redundancy of the diffusion process by connecting people whose friends do not know each other, thereby allowing a behavior to rapidly spread to other areas of the network (3–5). The ideal case for this lack of redundancy is a “random” network, in which, in expectation for a large population, each of an individual’s ties reaches out to different neighborhoods (4, 9). The other hypothesis states that, unlike disease, social behavior is a complex contagion: People usually require contact with multiple sources of “infection” before being convinced to adopt a behavior (2). This hypothesis predicts that because clustered networks have more redundant ties, which provide social reinforcement for adoption, they will better promote the diffusion of behaviors across large populations (2, 7). Despite the scientific (6, 7, 10) and practical (1, 2, 11) importance of understanding the spread of behavior

through social networks, an empirical test of these predictions has not been possible, because it requires the ability to independently vary the topological structure of a social network (12).

I tested the effects of network structure on diffusion using a controlled experimental approach. I studied the spread of a health behavior through a network-embedded population by creating an Internet-based health community, containing 1528 participants recruited from health-interest World Wide Web sites (13).

Each participant created an anonymous online profile, including an avatar, a user name, and a set of health interests. They were then matched with other participants in the study—referred to as “health buddies”—as members of an online health community. Participants could not contact their health buddies directly, but they could receive emails from the study informing them of their health buddies’ activities. To preserve anonymity and to prevent people from trying to identify

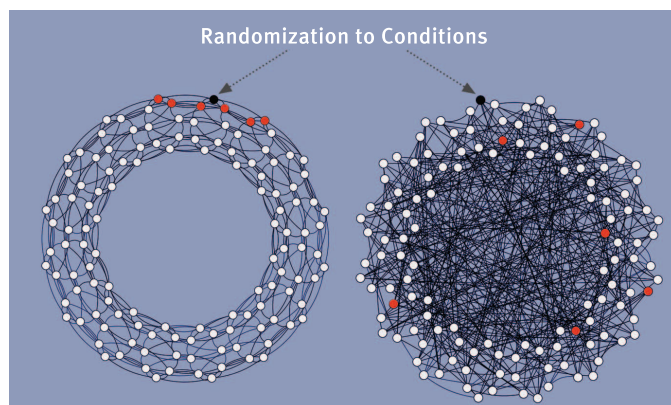
friends who may have also signed up for the study (or from trying to contact health buddies outside the context of the experiment), I blinded the identifiers that people used. Participants made decisions about whether or not to adopt a health behavior based on the adoption patterns of their health buddies. The health behavior used for this study was the decision to register for an Internet-based health forum, which offered access and rating tools for online health resources (13).

The health forum was not known (or accessible) to anyone except participants in the experiment. This ensured that the only sources of encouragement that participants had to join the forum were the signals that they received from their health buddies. The forum was populated with initial ratings to provide content for the early adopters. However, all subsequent content was contributed by the participants who joined the forum.

Participants arriving to the study were randomly assigned to one of two experimental conditions—a clustered-lattice network and a random network—that were distinguished only by the topological structure of the social networks (Fig. 1). In the clustered-lattice network condition, there was a high level of clustering (5, 6, 13) created by redundant ties that linked each node’s neighbors to one another. The random network condition was created by rewiring the clustered-lattice network via a permutation algorithm based on the small-world-network model (6, 13–15). This ensured that each node maintained the exact same number of neighbors as in the clustered network (that is, a homogeneous degree distribution), while simultaneously reducing clustering in the network and eliminating redundant ties within and between neighborhoods (4, 6, 14).

The network topologies were created before the participants arrived, and the participants could

**Fig. 1.** Randomization of participants to clustered-lattice and random-network conditions in a single trial of this study ( $N = 128$ ,  $Z = 6$ ). In each condition, the black node shows the focal node of a neighborhood to which an individual is being assigned, and the red nodes correspond to that individual’s neighbors in the network. In the clustered-lattice network, the red nodes share neighbors with each other, whereas in the random network they do not. White nodes indicate individuals who are not connected to the focal node.



Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02142, USA. E-mail: dcentola@mit.edu

not alter the topology in which they were embedded (e.g., by making new ties). In both conditions, each participant was randomly assigned to occupy a single node in one network. The occupants of the immediately adjacent nodes in the network (i.e., the network neighbors) constituted a participant's health buddies (13). Each node in a social network had an identical number of neighbors as the other nodes in the network, and participants could only see the immediate neighbors to whom they were connected.

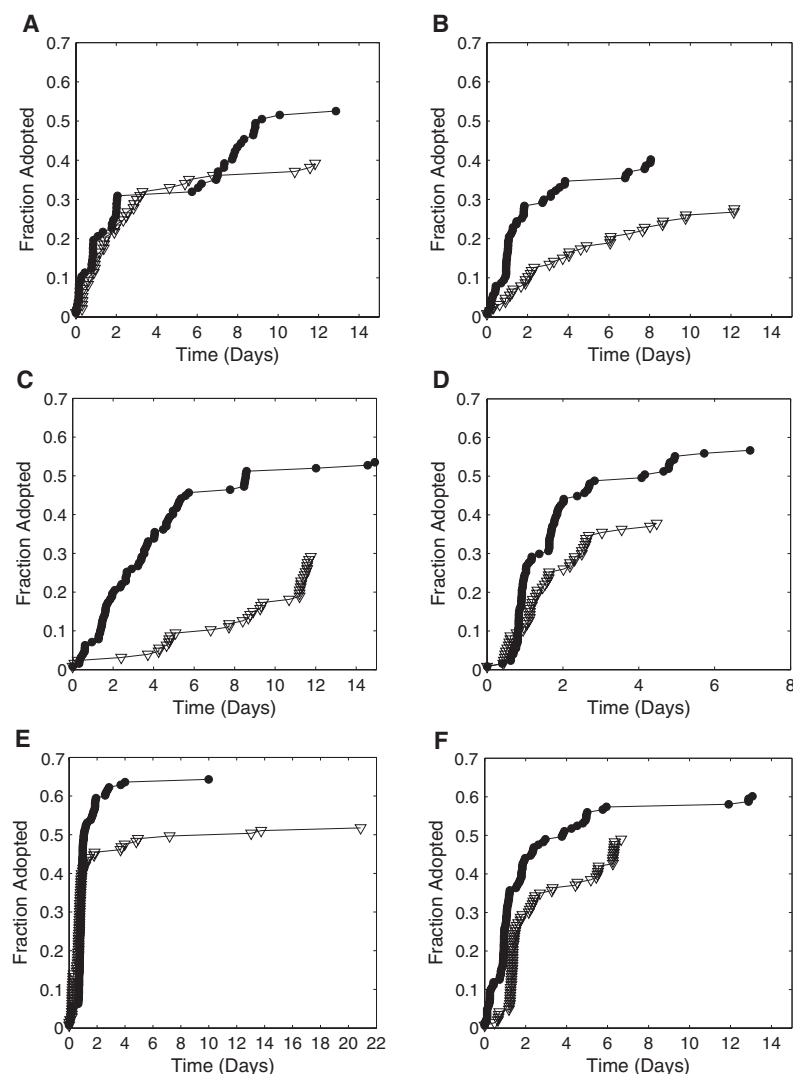
Consequently, the size of each participant's social neighborhood was identical for all participants within a network and across conditions. More generally, every aspect of a participant's experience before the initiation of the diffusion dynamics was equivalent across conditions, and the only difference between the conditions was the pattern of connectedness of the social net-

works in which the participants were embedded. Thus, any differences in the dynamics of diffusion between the two conditions can be attributed to the effects of network topology.

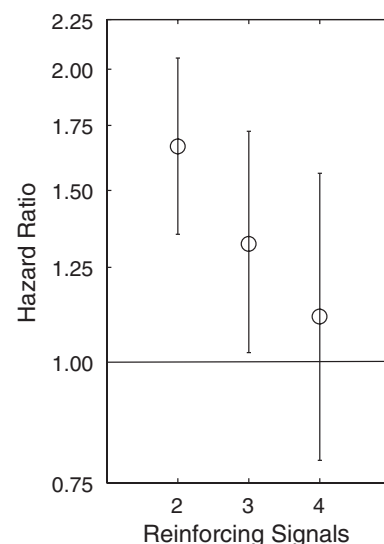
There are four advantages of this experimental design over observational data. (i) The present study isolates the effects of network topology, independent of frequently co-occurring factors such as homophily (3, 16), geographic proximity (17), and interpersonal affect (4, 18), which are easily conflated with the effects of topological structure in observational studies (2, 3, 11). (ii) I study the spread of a health-related behavior that is unknown to the participants before the study (13), thereby eliminating the effects of nonnetwork factors from the diffusion dynamics, such as advertising, availability, and pricing, which can confound the effects of topology on diffusion when, for example, the

local structure of a social network correlates with greater resources for learning about or adopting an innovation (11, 19). (iii) This study eliminates the possibility for social ties to change and thereby identifies the effects of network structure on the dynamics of diffusion without the confounding effects of homophilous tie formation (1, 20). (iv) Finally, this design allows the same diffusion process to be observed multiple times, under identical structural conditions, thus allowing the often stochastic process of individual adoption (21) to be studied in a way that provides robust evidence for the effects of network topology on the dynamics of diffusion.

I report the results from six independent trials of this experimental design, each consisting of a matched pair of network conditions. In each pair, participants were randomized to either a clustered-lattice network or a corresponding random network (13). This yielded 12 independent diffusion processes. Diffusion was initiated by selecting a random "seed node," which sent signals to its network neighbors encouraging them to adopt a health-related behavior—namely, registering for a health forum Web site (13). Every time a participant adopted the behavior (i.e., registered for the health forum), messages were sent to her health buddies inviting them to adopt. If a participant had multiple health buddies who adopted the behavior, then she would receive multiple signals, one from each neighbor. The more neighbors who adopted, the more reinforcing signals a participant received. The sequence of adoption decisions made by the



**Fig. 2.** Time series showing the adoption of a health behavior spreading through clustered-lattice (solid black circles) and random (open triangles) social networks. Six independent trials of the study are shown, including (A)  $N = 98$ ,  $Z = 6$ , (B to D)  $N = 128$ ,  $Z = 6$ , and (E and F)  $N = 144$ ,  $Z = 8$ . The success of diffusion was measured by the fraction of the total network that adopted the behavior. The speed of the diffusion process was evaluated by comparing the time required for the behavior to spread to the greatest fraction reached by both conditions in each trial.



**Fig. 3.** Hazard ratios for adoption for individuals receiving two, three, and four social signals. The hazard ratio  $g$  indicates that the likelihood of adoption increases by a factor of  $g$  for each additional signal  $k$ , compared to the likelihood of adoption from receiving  $k - 1$  signals. The 95% confidence intervals from the Cox proportional hazards model are shown by error bars. The effect of an additional signal on the likelihood of adoption is significant if the 95% confidence interval does not contain  $g = 1$  (13).

members of each social network provides a precise time series of the spread of the behavior through the population. It also provides an exact record of the number of signals required for individuals to adopt the behavior. The starting time (time = 0) for each diffusion process corresponds to the instant when the seed node was activated and the initial signals were sent. For each trial, the diffusion process was allowed to run for 3 weeks (~1.8 million seconds). To test for the possible effects of population size ( $N$ ) and degree ( $Z$ , the number of health buddies each person had) on the diffusion dynamics, I used three different versions of the experiment: (i)  $N = 98$ ,  $Z = 6$ ; (ii)  $N = 128$ ,  $Z = 6$ ; and (iii)  $N = 144$ ,  $Z = 8$  (13). The modest range of population sizes tested and the correspondingly narrow range of degrees were due to the challenges of recruiting large numbers of people simultaneously. Among the networks I used, there were no effects of population size (13).

The results show that network structure has a significant effect on the dynamics of behavioral diffusion. Surprisingly, the topologies with greater clustering and a larger diameter were much more effective for spreading behavior. Figure 2 shows the time series generated by the six independent trials of the experiment. Adoption typically spread to a greater fraction of the population in the clustered networks (solid black circles) than

in the random networks (open triangles). On average, the behavior reached 53.77% of the clustered networks, whereas only 38.26% of the population adopted in the random networks (13). I also found that the behavior diffused more quickly in the clustered networks than in the random networks. The average rate of diffusion in the clustered networks ( $0.2820 \times 10^{-3}$  nodes/s) was more than four times faster than that of the random condition ( $0.0643 \times 10^{-3}$  nodes/s). Differences in both the success and the rate of diffusion between network conditions are statistically significant ( $P < 0.01$  using the Wilcoxon rank sum–Mann–Whitney  $U$  test) (13).

The experimental findings were qualitatively the same across different network and neighborhood sizes. However, networks with a greater degree ( $Z = 8$ ) performed better than those with a lower degree ( $Z = 6$ ). Although this finding is consistent with the hypothesis that more redundant ties between neighborhoods can improve the global spread of behavior, it may also indicate that other topological features, such as degree and density, are relevant factors affecting behavioral diffusion (2, 7). This suggests important avenues for future research.

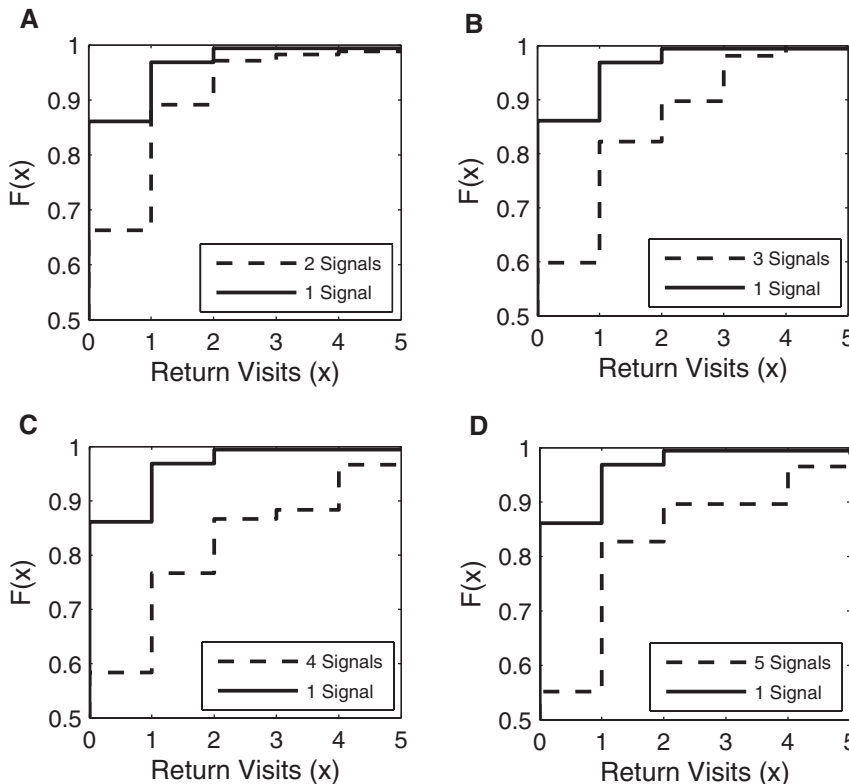
At the individual level, the results (Fig. 3) show that redundant signals significantly increased the likelihood of adoption; social rein-

forcement from multiple health buddies made participants much more willing to adopt the behavior. Figure 3 compares the baseline likelihood of adoption after receiving one social signal to the increased likelihood of adoption for nodes receiving second, third, and fourth reinforcing signals. Participants were significantly more likely to adopt after receiving a second signal than after receiving only one signal ( $P < 0.001$  using the Cox proportional hazards model). Receiving a third signal also significantly increased the likelihood of adoption, but with a smaller marginal-effect size ( $P < 0.05$ , Cox proportional hazards model) (13). Additional signals had no significant effect. This can be attributed to the attenuation of the sample size as the number of signals increased.

A secondary, but important, issue related to adoption is the level of commitment that individuals have to a behavior once they have adopted it. To investigate the effects of social reinforcement on individuals' level of engagement with the health forum, I compared the number of return visits to the forum after registering, for adopters grouped by the number of social signals that they received (Fig. 4) (participants could not receive additional signals once they had registered). Figure 4 shows pairwise comparisons of the number of return visits for adopters receiving only one signal (solid lines) versus those receiving two to five signals (dashed lines in panels A to D, respectively). Though less than 15% of adopters receiving one signal made a return visit to the forum, more than 30% of participants receiving two signals made return visits, and 40% of participants receiving three signals made at least one return visit. Pairwise statistical comparisons between group one and groups two through five are all significant ( $P < 0.01$  for all four comparisons, using the Kolmogorov–Smirnov test) (13), indicating that participants who received more than one social signal were significantly more likely to return to the health forum than those who only received a single signal. This suggests that there was a significant effect of social reinforcement on participants' level of engagement with the adopted behavior.

As with all experiments, design choices that aided my control of the study also put constraints on the behaviors that I could test. A key limitation of my design is that, unlike in my experiment, adopting a new health behavior is often extremely difficult in the real world. To adopt behaviors such as getting a vaccination, going on a diet, starting an exercise routine, or getting a screening, people may be required to pay the costs of time, deprivation, or even physical pain. Because of this, I expect that the need for social reinforcement would be greater for adopting these health behaviors than it was for the behavior in my study. Consequently, the diffusion of real-world health behaviors may depend even more on clustered-network structures than did the diffusion dynamics reported in my results.

An additional constraint of my study was that participants did not have any direct commu-



**Fig. 4.** Cumulative distribution functions of the number of return visits to the health forum ( $x$ ) for populations of adopters grouped by the number of signals that they received. Comparisons are shown for adopters who received (A) one versus two signals, (B) one versus three signals, (C) one versus four signals, and (D) one versus five signals. All pairwise comparisons between groups two through five with each other showed no significant differences ( $P > 0.4$  for all six comparisons, using the Kolmogorov–Smirnov test) (13).



nication with their health buddies or information about their identities. This allowed me to isolate the effects of network topology on the dynamics of diffusion without the presence of confounding variables. However, it also raises the question of what the strength of the effects of network topology would be when allowed to interact with the effects of interpersonal relationships. An important assumption of this study is that the effects of network topology will not be overwhelmed by individuals' exposure to other social factors. Previous studies have suggested that factors such as homophily and strong interpersonal affect in social ties can improve the diffusion of behaviors through social networks (3, 18). In the real world, these features of social relationships tend to be highly correlated with the formation of clustered social ties (3, 22, 23). Consequently, I expect that these reinforcing factors would amplify the observed effects of clustered social networks in promoting the diffusion of health behaviors across a large population. However, new experimental designs are required to test the interaction effects of these variables (and other variables such as gender, memory, and frequency of interaction) on the spread of social behaviors.

Evidence in support of the "strength of weak ties" hypothesis has suggested that networks with high levels of local clustering and tightly knit neighborhoods are inefficient for large-scale diffusion processes (4, 5, 9). My findings show that, not only is individual adoption improved by reinforcing signals that come from clustered social ties (Fig. 3), but this individual-level effect also translates into a system-level phenomenon where by large-scale diffusion can reach more people

and spread more quickly in clustered networks than in random networks (Fig. 2). Whereas locally clustered ties may be redundant for simple contagions, like information or disease (4, 6, 24), they can be highly efficient for promoting behavioral diffusion. On the basis of these findings, I predict that public health interventions aimed at the spread of new health behaviors (for instance, improved diet, regular exercise, condom use, or needle exchange) may do better to target clustered residential networks rather than the casual contact networks across which disease may spread very quickly (25)—particularly if the behaviors to be diffused are highly complex (for instance, because they are costly, difficult, or contravene existing norms).

#### References and Notes

- N. A. Christakis, J. H. Fowler, *N. Engl. J. Med.* **357**, 370 (2007).
- D. Centola, M. Macy, *Am. J. Sociol.* **113**, 702 (2007).
- E. M. Rogers, *Diffusion of Innovations* (Free Press, New York, 1995).
- M. Granovetter, *Am. J. Sociol.* **78**, 1360 (1973).
- D. J. Watts, *Small Worlds: The Dynamics of Networks Between Order and Randomness* (Princeton Univ. Press, Princeton, NJ, 1999).
- D. J. Watts, S. H. Strogatz, *Nature* **393**, 440 (1998).
- D. Centola, V. Eguiluz, M. Macy, *Phys. A* **374**, 449 (2007).
- N. A. Christakis, J. H. Fowler, *N. Engl. J. Med.* **358**, 2249 (2008).
- M. E. J. Newman, *J. Stat. Phys.* **101**, 819 (2000).
- S. H. Strogatz, *Nature* **410**, 268 (2001).
- K. P. Smith, N. A. Christakis, *Annu. Rev. Sociol.* **34**, 405 (2008).
- M. Kearns, S. Suri, N. Montfort, *Science* **313**, 824 (2006).
- Materials and methods are available as supporting material on Science Online.

- S. Maslov, K. Sneppen, *Science* **296**, 910 (2002).
- S. Maslov, K. Sneppen, U. Alon, in *Handbook of Graphs and Networks*, S. Bornholdt, H. G. Schuster, Eds. (Wiley-VCH, Weinheim, Germany, 2003).
- J. M. McPherson, L. Smith-Lovin, J. M. Cook, *Annu. Rev. Sociol.* **27**, 415 (2001).
- P. Hedstrom, *Am. J. Sociol.* **99**, 1157 (1994).
- D. McAdam, R. Paulsen, *Am. J. Sociol.* **99**, 640 (1993).
- L. F. Berkman, I. Kawachi, *Social Epidemiology* (Oxford Univ. Press, New York, 2000).
- D. Centola, J. C. Gonzalez-Avella, V. Eguiluz, M. San Miguel, *J. Conf. Resl.* **51**, 905 (2007).
- M. J. Salganik, P. S. Dodds, D. J. Watts, *Science* **311**, 854 (2006).
- G. Kossinets, D. J. Watts, *Am. J. Sociol.* **115**, 405 (2009).
- G. Kossinets, D. J. Watts, *Science* **311**, 88 (2006).
- M. E. J. Newman, D. J. Watts, *Phys. Rev. E* **60**, 7332 (1999).
- P. D. Calvert *et al.*, *Nature* **411**, 90 (2001).
- I thank A. van de Rijt and V. Eguiluz for helpful comments, N. Christakis for useful discussion and guidance, G. Pickard for programming assistance, A. Wagner for developing the Healthy Lifestyle Network Web site, T. Groves for all of the design work, and J. Kreckler and K. Campbell at [www.prevention.com](http://www.prevention.com) and G. Colditz of [www.yourdiseaserisk.com](http://www.yourdiseaserisk.com) for their assistance recruiting participants. This work was supported in part by the James S. McDonnell Foundation and the Robert Wood Johnson Foundation.

#### Supporting Online Material

[www.sciencemag.org/cgi/content/full/329/5996/1194/DC1](http://www.sciencemag.org/cgi/content/full/329/5996/1194/DC1)

Materials and Methods

SOM Text

Figs. S1 to S7

Tables S1 to S3

References

26 November 2009; accepted 19 July 2010  
10.1126/science.1185231

## Human-Restricted Bacterial Pathogens Block Shedding of Epithelial Cells by Stimulating Integrin Activation

Petra Muenzner,<sup>1</sup> Verena Bachmann,<sup>1</sup> Wolfgang Zimmermann,<sup>2</sup>  
Jochen Hentschel,<sup>1,3</sup> Christof R. Hauck<sup>1,4,\*</sup>

Colonization of mucosal surfaces is the key initial step in most bacterial infections. One mechanism protecting the mucosa is the rapid shedding of epithelial cells, also termed exfoliation, but it is unclear how pathogens counteract this process. We found that carcinoembryonic antigen (CEA)-binding bacteria colonized the urogenital tract of CEA transgenic mice, but not of wild-type mice, by suppressing exfoliation of mucosal cells. CEA binding triggered de novo expression of the transforming growth factor receptor CD105, changing focal adhesion composition and activating  $\beta_1$  integrins. This manipulation of integrin inside-out signaling promotes efficient mucosal colonization and represents a potential target to prevent or cure bacterial infections.

During colonization of mucosal surfaces, incoming microbes must cope with multiple host defenses (1). One protective mechanism of the mucosa in stratified and squamous tissues is the accelerated turnover and shedding of superficial epithelial cells, also referred to as exfoliation (2–4). Although in vitro

studies have suggested that microbes modulate cell detachment (5, 6), it is currently unknown how successful mucosal pathogens deal with the exfoliation response in vivo. *Neisseria gonorrhoeae*, a Gram-negative microorganism, causes one of the most common sexually transmitted diseases worldwide (7). Even though these bacteria can

induce the exfoliation of host cells upon contact (8–11), they are able to establish themselves on virtually every mucosal surface of the human body.

To investigate how gonococci manage to colonize the urogenital mucosa efficiently, we performed vaginal infection of female mice (12). In line with the innate capacity of epithelial cells to respond to this bacterial challenge, *N. gonorrhoeae* triggered detachment of superficial epithelial cells within 20 hours (Fig. 1A) and only small numbers of gonococci could be re-isolated from wild-type mice (Fig. 1B). Gonococci are adapted to humans as their sole natural host. One of the host-specific virulence traits that gonococci share with other specialized mucosal colonizers, including *Haemophilus influenzae*, *Moraxella catarrhalis*, and *N. meningitidis*, is the ability to recognize hu-

<sup>1</sup>Lehrstuhl Zellbiologie, Fachbereich Biologie, Universität Konstanz, 78457 Konstanz, Germany. <sup>2</sup>Tumor Immunology Laboratory, LIFE Center, Ludwig-Maximilians-Universität München, 81377 München, Germany. <sup>3</sup>EM Service, Fachbereich Biologie, Universität Konstanz, 78457 Konstanz, Germany. <sup>4</sup>Konstanz Research School Chemical Biology, Universität Konstanz, 78457 Konstanz, Germany.

\*To whom correspondence should be addressed. E-mail: [christof.hauck@uni-konstanz.de](mailto:christof.hauck@uni-konstanz.de)