Spatio-temporal propagation of cascading overload failures in spatially embedded networks

Jichang Zhao1, Daqing Li2,3, Hillel Sanhedrai4, Reuven Cohen5 & Shlomo Havlin4

Different from the direct contact in epidemics spread, overload failures propagate through hidden functional dependencies. Many studies focused on the critical conditions and catastrophic consequences of cascading failures. However, to understand the network vulnerability and mitigate the cascading overload failures, the knowledge of how the failures propagate in time and space is essential but still missing. Here we study the spatio-temporal propagation behaviour of cascading overload failures analytically and numerically on spatially embedded networks. The cascading overload failures are found to spread radially from the centre of the initial failure with an approximately constant velocity. The propagation velocity decreases with increasing tolerance, and can be well predicted by our theoretical framework with one single correction for all the tolerance values. This propagation velocity is found similar in various model networks and real network structures. Our findings may help to predict the dynamics of cascading overload failures in realistic systems.
Resilience of individual components in networks is determined not only by their own intrinsic properties, but also by their functional interactions with other components. For example, a failure of one component in a network may lead to overloads and failures of other components. Starting with a localized failure, such interactions between components can ignite a domino-like cascading failure, which may result in catastrophes such as those observed in many realistic networks. The devastating consequences of these cascading failures have stimulated extensive studies.

Many studies provided deep insight on the conditions and outcome of cascading failures, which may help to evaluate the system resilience. However, to predict and mitigate the failure spreading in a network, understanding their spatio-temporal propagation properties is essential but still missing. Different from structural cascading failures caused by direct causal dependencies, overload failures usually propagate through invisible paths as a result of cooperative interactions in the system. Actually, a fundamental question has rarely been posed: how overload failures propagate with time in space during the cascading process? Indeed, predicting the spatio-temporal propagation of cascading failures could determine the timing and resource allocation of an effective mitigation strategy in corresponding self-healing technologies.

In this paper, we aim to understand the spatio-temporal propagation of the cascading overload failures in spatially embedded systems. We define two basic quantities to describe the spatio-temporal propagation properties of cascading failures. The first quantity, \( r_c(t) \), is the average Euclidean distance of the failures appearing at cascading step \( t \) from the centre of the initial failure. The second quantity, \( F_c(t) \), is the number of node failures that occur at cascading step \( t \). While \( r_c \) can help system regulators to set a ‘firewall’ at suitable locations before the failure arrives, \( F_c \) can suggest the ‘height’ of the firewall. In our current study, the propagation of cascading overloads is found to follow an approximately constant velocity. This propagation velocity decreases as the system tolerance increases, which can be well predicted by our theoretical analysis with one single constant correction. The propagation velocity is also found here to be similar in various model networks and real network structures.

**Results**

**Propagation properties of cascading overload failures.** In this study, we focus on the cascading overloads on spatially embedded networks caused by localized attacks, which are common in natural disaster and malicious attacks. In Fig. 1, we show snapshots of the simulated cascading failures, where the failures spread almost radially from the initial attack region and finally spread over the whole system. As can be seen from Fig. 1, the nodes with spatial location closer to the initial failure begin to fail first and form approximately a ‘ring of failures’. The ring begins to grow and expand with time until it reaches the system’s boundary.

Considering the ring shape of the cascading overload failures (originating from the initial location), it is reasonable to quantify the propagation radius of failures, \( r_c(t) \), and the number of failures, \( F_c(t) \) at cascading step \( t \). Figure 2a shows that the propagation radius of cascading failures is increasing almost linearly with time for different system sizes. This means that the failures spread during the cascade process with an almost constant speed (slope of \( r_c(t) \)). It can be seen that as the system linear size \( L \) increases, the propagation velocity of overload failures increases. The propagation radius of failures is increasing with time and becomes saturated near the boundary. The propagation size of failure \( F_c \) (number of new failures at each instant) increases until a certain time step and then decreases (Fig. 2b). The behaviour at large \( t \) is due to the finite size of the system, when \( r_c(t) \) reaches the order of \( L \), the amount of damage can only decrease. Note that for different system sizes, the propagation pattern of failure \( F_c \) reaches the maximum at similar instants, which results from the higher velocity in larger system sizes.

To further understand the effect of system size and tolerance \( \alpha \) on failure propagation \( r_c(t) \), we rescale it in Fig. 2c by the system linear size \( L \). We find that the curves of rescaled \( r_c(t) \) collapse into a single curve for different system sizes at a given tolerance \( \alpha \), suggesting that failures spread in the same relative velocities for different system sizes. When the size of failures, \( F_c(t) \), is rescaled by the system size, \( L \times L \), the different curves of \( F_c(t)/L^2 \) also collapse into a single curve (Fig. 2d). It can be seen that the spreading time needed to reach the maximum \( F_c(t)/L^2 \) is determined by the tolerance \( \alpha \), which implies that large tolerance can postpone the collapse of the system.

In contrast to the dynamics such as epidemics that propagate owing to nearest-neighbor interactions such as contact infection, cascading overloads propagate as a result of the global interaction between all the flows contributed by the whole system. Surprisingly, although the interactions are global, the propagation dynamics in the model network and realistic networks (see Supplementary Figs 1–6 and Supplementary Notes 1 and 2 for more examples) are found rather local. Here the local propagation means that there is a finite characteristic distance \( \langle A(r_c) \rangle \) between the successive overloads. This characteristic distance is the value of propagation velocity, which increases non-linearly with decreasing tolerance. The nearest-neighbor propagation of overloads usually assumed in some complex

---

**Figure 1 | The propagation of the overload failures in the network.** We demonstrate step 1, 3, 5 and 7 of the cascading failures on a 200 × 200 lattice with periodic boundary conditions and a Gaussian distribution of weights. The disorder is \( \alpha = 0.01 \), the initial attack size (in red) is 6 × 6 and the tolerance of system is set to \( \alpha = 0.5 \). In each figure, the deep blue dots stands for the overloaded nodes in the current step, while the black ones are the nodes failed in the previous steps. The cyan dots are the functional nodes that did not fail. (a) \( t = 1 \), (b) \( t = 3 \), (c) \( t = 5 \) and (d) \( t = 7 \).
Simulations are limited by the computational complexity of the most efficient algorithm for calculating node betweenness, which is \( O(NM + N^2 \log N) \) for weighted networks, where \( N \) is the system size and \( M \) is the number of edges (\( NM \) is the order of \( N^2 \) for sparse network).

Propagation results based on theoretical analysis. To further explore these propagation behaviours of cascading failures found in simulations and their relations with tolerance, we develop the theory (see Fig. 3, Methods section and Supplementary Methods for more details) to describe the cascading overloads.

As can be seen in Fig. 4, our theoretical analysis reproduces well the spatio-temporal propagation features of cascading overloads found in the simulations (Fig. 2). Specifically, as shown in Fig. 4a,c, the velocity of the failure propagation is almost constant in most of the time \( t \), and decreases with increasing tolerance or decreasing system size. As shown in Fig. 4b,d, the number of failures increases with time and then begins to drop after reaching the peak, which is reduced by increasing the tolerance \( \alpha \). The instant for the maximal failure size is independent of the system size, but increases with the system tolerance. Moreover, similar to simulation results, both the radius and size of failures in the cascades can also be rescaled with \( \alpha \) and \( \Delta \) are the results of \( r_c(\tau) \) and \( F(\tau) \) scaled by the system size, including \( L = 70 \) (triangle), \( L = 80 \) (square), \( L = 90 \) (circle) and \( L = 100 \) (diamond), respectively. The results are obtained by averaging over fifty realizations. Note that simulations are limited by the computational complexity of the most efficient algorithm for calculating node betweenness, which is \( O(NM + N^2 \log N) \) for weighted networks, where \( N \) is the system size and \( M \) is the number of edges (\( NM \) is the order of \( N^2 \) for sparse network).

Comparison between simulation and theory. To test our theory, we perform a quantitative comparison between the simulation and the theoretical model. As shown in Fig. 5a, the propagation velocity in different model networks and real networks can be well predicted by the theory with the same constant correction for

**Figure 2 | Spatio-temporal propagation of cascading overload failures in simulations.** (a,b) are the spreading radius \( r_c(\tau) \) of failures and amount of failures at each step, \( F(\tau) \), as a function of time for \( \alpha = 0.25, \sigma = 0.1 \) and \( L = 6 \). (c,d) is the number of failures and amount of failures and \( F(\tau) \) scaled by the system size, including \( L = 70 \) (triangle), \( L = 80 \) (square), \( L = 90 \) (circle) and \( L = 100 \) (diamond), respectively. The results are obtained by averaging over fifty realizations. Note that simulations are limited by the computational complexity of the most efficient algorithm for calculating node betweenness, which is \( O(NM + N^2 \log N) \) for weighted networks, where \( N \) is the system size and \( M \) is the number of edges (\( NM \) is the order of \( N^2 \) for sparse network).

**Figure 3 | Theory for overload propagation.** The network is embedded in a two-dimensional circular plate centred at \( O \) with a radius of 1 unit and the initial failure is located at the centre of the network within a circle of radius \( a \approx 1 \). The ring centred at \( O \) and between \( a \) and \( b \) (\( b > a \)) is defined as the adjacent ring. \( A \) is a random node in the network, whose distance to \( O \) is \( r \). The ring \( B \) is the original path starting from \( A \) to a random node \( F \) on the system border. Since \( r > b \), the intersection points between \( AF \) and the circle with radius \( b \) are \( B \) and \( E \). \( AF \) also intersects the border of the failure area at \( C \) and \( D \). \( AD \) is a straight line tangent to the failure area and the tangent point is \( G \). We also define another path \( AI \), which starts from \( A \) to a border node \( I \) by passing through the plate centre \( O \). Note that a realistic road network is embedded behind the circular plate for demonstration.
all the values of tolerances. This constant correction, close to $2\pi$, is a result of the anisotropic propagation of overloads in the simulation, which is assumed isotropic in the theory. These good agreements between theory and simulations support the validity of our proposed theoretical framework for cascading overloads.

The local propagation found in the simulation and theory is due to the mechanism of overload redistribution. In the simulation of the Motter–Lai model in Fig. 5b, the overloads propagate rather locally (with a characteristic distance) after an initial damage due to the redistribution of optimal paths between existing pairs of nodes. The optimal paths that passed through the previous damaged area will now mainly be redistributed close to this area, leading to the local overloads in spite of global interactions. Similar to our findings in the Motter–Lai model, the overloads in the theory (as shown in Fig. 5b and Supplementary Fig. 10) are found here mainly distributed close to the previous damage area, which causes the local propagation of failures (also as shown in Supplementary Fig. 11 and Supplementary Note 4).

Note the agreement in Fig. 5b between theory and simulations (with the same constant correction of $2\pi$, due to the anisotropic...
nature of the failure spreading, which is assumed isotropic in the theory). The excellent agreement between simulations and theory of the overload (Fig. 5b) as a function of distance from the original failure also supports our theoretical approach.

It is worth noting that the analysis of our theoretical framework is not limited by the computational complexity of calculating optimal paths. This is in contrast to simulations based on betweenness (number of optimal paths passing through a node), which require heavy computations and therefore are limited to relatively small systems. Our theoretical framework thus provides an efficient way to explore and understand the cascading behaviour of failures for any system size.

**Universality of propagation velocity.** To explore the possibility of universal feature in overload propagation, here we analyse the propagation of cascading overload failures on different network models and realistic networks. Besides the model networks including lattice, circular lattice, planar graph and circular plate, we also study the overload propagation on the realistic road networks. We initiate a local attack in the geographical centre of these spatially embedded networks and analyse the spatio-temporal evolution of $r$ and $F$. As shown in Fig. 5a (see Supplementary Figs 1–6 for more examples), we find that propagation velocity of overloads is similar in various model and real network structures, which can be well predicted by our theory. The propagation velocity is independent of detailed network structures, which makes our findings more applicable. This universal propagation of overloads can be attributed to the common mechanism of overload redistribution in different networks, independent of structure difference between these networks. Furthermore, both theory and simulations suggest that for a given system size and a given tolerance, the size of initial failure does not influence the spreading velocity (Supplementary Figs 12 and 13 and Supplementary Notes 5 and 6).

**Discussion**

Cascading failures represent the manifestation of non-linear butterfly effect in infrastructure networks, which can cause catastrophic damages due to a small local disturbance. The Motter–Lai-type overload cascade models are an important class of cascading failure dynamics, characterized by the nonlocal—in contrast to epidemic spreading-type local cascade models—interactions, and have been studied extensively for the last decade. Given the inherent global interactions in the mechanism of overload formation, it is of interest that the overloads spread rather locally from the initial attack region, at velocity that increases non-linearly with decreasing tolerance. Here the local propagation means that there is a finite characteristic distance between the successive overloads, which is the value of propagation velocity. The nearest-neighboring propagation of overloads usually assumed in some complex network models only corresponds to the limiting case of this local propagation (characteristic distance is 1).

For different model and realistic networks, our results suggest the existence of universal propagation features of cascading overloads, which are characterized by a finite linear propagation velocity. This velocity can be predicted by our theory with the same constant correction for all the values of tolerances. This constant correction, close to $2\pi$, is a result of the anisotropic propagation of overloads in the simulation, which is assumed isotropic in the theory. This universal behaviour comes from the common global mechanisms of overload redistributions in different networks. When certain extreme heterogeneous networks like embedded scale free networks are considered, a revision of the theoretical framework may be needed. While the focus of this manuscript is on the spatio-temporal propagation of the cascading overload failures in spatially embedded systems, the overloads may spread very fast in general non-spatially embedded sparse network due to its small diameter.

The present study may help to bridge the longstanding gap between the overload model 16,22 and the model of dependency links proposed by Buldyrev et al. 13 and Parshani et al. 19, in particular the lattice version of the model 16,20 where dependency links can have a characteristic length $r$. Indeed, as can be seen in Figs 2 and 4, overload failures propagate in a nearly constant speed, which suggests a characteristic dependency distance, between successive overload failures. Furthermore, this speed or characteristic distance is found to increase with decreasing tolerance. This suggests a possible mapping between systems with overload failures and networks with dependency links, where networks with different characteristic length of dependency links can serve as a suitable model to describe cascading overload failures. This mapping can be useful since overload models usually require heavy computations and are therefore limited to small systems, while dependency models require significantly less computations, and larger systems can be easily analysed.

When a disturbance is detected in networks, the knowledge of spatio-temporal propagation properties of cascading failures is essential for predicting and mitigating the cascading failures. Meanwhile, realistic cascading failures are usually the result of the collective interactions between different processes including overloads and other system operation procedures 30–32. The universal features of overload propagation found here across different spatially embedded networks may help to better mitigate realistic cascading failure, if combined with the detailed knowledge of other processes including system operations and planning procedures.

**Methods**

**Simulation of model.** To study the propagation properties of cascading overload failures, we model the spatially embedded network as a randomly weighted $L \times L$ lattice with periodic boundaries, where $L$ is the linear length of the lattice. The weight of each link is taken from a Gaussian distribution $N(\mu, \sigma^2)$ with mean weight $\mu$ and the disorder is represented by the standard deviation $\sigma$. In this model 22, the load on a node $L_i$, represents the number of optimal paths between all pairs of nodes passing through this node. A node will fail when its load $L_i$ is more than $(1 + z)$ times its original load, where $z$ represents the system tolerance to overloads. A randomly localized region of the system is initially removed to trigger the cascading overloads. This kind of initial failure is motivated by the fact that natural disasters (like earthquake or floods) or malicious attacks (like electromagnetic pulse (EMP)) usually occur in specific geographical locations and destroy initially localized regions of the network. The initial failure is located in a randomly selected square of $I \times I$ nodes ($I < \sqrt{L}$), which are removed initially. In addition to being realistic, since failures are usually localized, this configuration can also help us to follow and analyse the spatial-temporal propagation pathway of the cascading failures. This local failure may trigger failures of other nodes if their load value exceeds the tolerance threshold due to the load redistribution across the entire network. Given the periodic boundary conditions, we can position the initial attack region at the centre of the lattice.

**Theoretical analysis.**

Cascading overload failures due to an initial local failure are produced by the redistribution of loads in the network. From the observations of simulations on weighted network, we find (Fig. 1) that failures spread in a ring shape from the centre of the initial damage. This inspires us to assume in our theory that the network is embedded in a two-dimensional circular plate (Fig. 3), the initial failure is within a (red) circle of radius $a$, and the main overload due to traffic after the initial failure is located in the (cyan) ring adjacent to the initial failure (see the cyan ring in Fig. 3), whose size will be determined by the theory. The overload is reflected by the increase of the number and lengths of shortest paths passing through this ring. If the overload exceeds the capacity tolerance of a node $1 + z$ times its original load) within the adjacent ring, the node will fail in the next step, causing the cascading failure to propagate forward. Owing to the initial failed area, shortest paths from a given node $A$ to destinations located in the shadow area $S$ (the dotted area in Fig. 3) would be affected and become longer. Similarly, nodes have to surround the longest paths from $A$ to nodes in $S$ (for example, $AF$ across the failed area can be separated into two parts in the adjacent ring (for example, $BC$ and $DE$). For the first part, its
length within the ring changes from BC to KG. As for the second part, its overload on the ring can be calculated from symmetry by switching A and F (source and target). Finally, the integration from \( r = a \) to \( r = 1 \) covers all the overlodes added to the adjacent ring (cyan) due to the initial failure, can be written as

\[
\Delta I_r(a,b) = \int_a^1 \left[ \sqrt{r^2 - a^2} s(a,r) - v(a,b,r)s(a,r) + \pi a^2 r \right] 2\pi r dr \]

\[
+ \int_a^1 \left[ \sqrt{b^2 - a^2} s(a,r) - v(a,b,r)s(a,r) + \pi a^2 r \right] 2\pi r dr,
\]

where the length of KG is \( \sqrt{b^2 - a^2} \) for \( r > b \) and \( \sqrt{r^2 - a^2} \) for \( r < b \). \( a, x, r \) is the area of a, that is, the number of node's destinations in the shadow and \( v(a,b,r) \) is the average length of the first part of shortest paths (within the adjacent ring) before the failure. Similarly, we can obtain the initial load of nodes located within the circle at \( O \) with radius \( b \) as \( I_{0_L}(b) \) (Supplementary Methods). Then, for a node on the circle centered \( O \) with radius \( b \), the overload produced by the failure is

\[
\Delta I_r = \frac{1}{2\pi b} \Delta I_r(a,b)
\]

and its initial load is

\[
I_o = \frac{1}{2\pi b} I_{0_L}(b)
\]

For each functional node in the network, the critical condition for failure can be written as \( \Delta = \Delta L/I_o \). Specifically, if \( \Delta \geq \Delta L/I_o \), it survives, otherwise it fails. More details for the solution of the theoretical model can be found in Supplementary Figs 14–19 and Supplementary Methods.

References

1. Helbing, D. Globally networked risks and how to respond. Nature 497, 51–59 (2013).
2. Dobson, L., Carreras, B. A., Lynch, V. E. & Newman, D. E. Complex systems analysis of series of blackouts: cascading failure, critical points, and self-organization. Chaos 17, 026103 (2007).
3. Toroczkai, Z. & Bassler, K. E. Network dynamics: jamming is limited in scale-free systems. Nature 428, 716 (2004).
4. Caldarelli, G., Chessa, A., Pammolli, F., Gabrielli, A. & Puliga, M. Reconstructing a credit network. Nat. Phys. 9, 125–126 (2013).
5. Frank, K. T., Petri, B., Choi, J. S. & Leggett, W. C. Topological cascades in a formerly co-dominant ecosystem. Science 327, 1621–1623 (2005).
6. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. Rev. Mod. Phys. 74, 47–97 (2002).
7. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. & Hwang, D.-U. Complex networks: Structure and dynamics. Phys. Rep. 424, 175–308 (2006).
8. Barrat, A., Barthelemy, M. & Vespignani, A. Dynamical Processes on Complex Networks (Cambridge University Press, 2008).
9. Cohen, R. & Havlin, S. Complex Networks: Structure, Robustness and Function (Cambridge University Press, 2010).
10. Watts, D. J. A simple model of global cascades on random networks. Proc. Natl Acad. Sci. 99, 5766–5771 (2002).
11. Lorentz, J., Battiston, S. & Schweitzer, F. Systemic risk in a unifying framework for cascading processes on networks. Eur. Phys. J. B 71, 441–460 (2009).
12. Araújo, N. A., Andrade, J. S., Ziff, R. M. & Herrmann, H. J. Tricritical point in explosive percolation. Phys. Rev. Lett. 106, 095703 (2011).
13. Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E. & Havlin, S. Catastrophic cascade of failures in interdependent networks. Nature 464, 1025–1028 (2010).
14. Zhao, J. H., Zhou, H. J. & Liu, Y. Y. Inducing effect on the percolation transition in complex networks. Nat. Commun. 4, 2412 (2013).
15. Radicchi, F. & Arenas, A. Abrupt transition in the structural formation of interconnected networks. Nat. Phys. 9, 717–720 (2013).
16. Motter, A. E. Cascade control and defense in complex networks. Phys. Rev. Lett. 93, 098701 (2004).
17. Simonson, I., Buzna, L., Peters, K., Bornholdt, S. & Helbing, D. Transient dynamics increasing network vulnerability to cascading failures. Phys. Rev. Lett. 100, 218701 (2008).
18. Parshani, R., Buldyrev, S. V. & Havlin, S. Critical effect of dependency groups on the function of networks. Proc. Natl Acad. Sci. USA 108, 1007–1010 (2011).
19. Li, W., Bashan, A., Buldyrev, S. V., Stanley, H. E. & Havlin, S. Cascading failures in interdependent lattice networks: the critical role of the length of dependency links. Phys. Rev. Lett. 108, 218701 (2011).
20. Bashan, A., Berezin, Y., Buldyrev, S. V. & Havlin, S. The extreme vulnerability of interdependent spatially embedded networks. Nat. Phys. 9, 667–672 (2013).
21. Goh, K.-I., Kahng, B. & Kim, D. Universal behavior of load distribution in scale-free networks. Phys. Rev. Lett. 87, 278701 (2001).
22. Motter, A. E. & Lai, Y. C. Cascade-based attacks on complex networks. Phys. Rev. E 66, 065102 (2002).
23. Crucitti, P., Latora, V. & Marchiori, M. Model for cascading failures in complex networks. Phys. Rev. E 69, 045104 (2004).
24. Li, D., Jiang, Y., Kang, R. & Havlin, S. Spatial correlation analysis of cascading failures: congestions and blackouts. Sci. Rep. 4, 5381 (2014).
25. Li, D., Kosmidis, K., Bunde, A. & Havlin, S. Dimension of spatially embedded networks. Nat. Phys. 7, 481–484 (2011).
26. Neumayer, S., Zussman, G., Cohen, R. & Modiano, E. Assessing the vulnerability of the fiber infrastructure to disasters. IEEE/ACM Trans. Netw. 19, 1610–1623 (2011).
27. Bernstein, A., Biesbrock, D., Hay, D., Usunugula, M. & Zussman, G. Power Grid Vulnerability to Geographically Correlated Failures - Analysis and Control Implications (Columbia University, Electrical Engineering, Technical Report, 2011).
28. Pastor-Satorras, R. & Vespignani, A. Epidemic spreading in scale-free networks. Phys. Rev. Lett. 86, 3200 (2001).
29. Li, F. Real Datasets for Spatial Databases: Road Networks and Points of Interest. http://www.cs.utah.edu/~lifeoi/SpatialDataset.htm (2013).
30. Menck, P. J., Heitzig, J., Kurths, J. & Schellnhuber, H. J. How dead ends undermine power grid stability. Nat. Commun. 5, 3969 (2014).
31. Cornelius, H. J., Kath, W. L. & Motter, A. E. Realistic control of network dynamics. Nat. Commun. 4, 1942 (2013).
32. Motter, J. E., Thorp, J. S. & Dobson, I. Cascading dynamics and mitigation assessment in power system disturbances via a hidden failure model. Int. J. Electr. Power Energy Syst. 27, 318–326 (2005).

Acknowledgements

We thank the support from Collaborative Innovation Center for industrial Cyber-Physical System. D.L. is also supported by the National Natural Science Foundation of China (Grant 61104444) and the National Basic Research Program of China (2012CB725040). S.H. thanks DTRA, ONR, the LINC and the Multiplex (No. 517532) EU projects, the DFG, and the Israel Science Foundation for support. J.Z. was partially supported by NSFC (Grant Nos 71501005 and 71531001) and 863 Program (Grant No. SS2014AA012303).

Author contributions

All authors contributed to all aspects of this work.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/