LETTERS

The abundance threshold for plague as a critical percolation phenomenon

S. Davis[1], P. Trapman[2], H. Leirs[3,4], M. Begon[5] & J. A. P. Heesterbeek[1]

Percolation theory is most commonly associated with the slow flow of liquid through a porous medium, with applications to the physical sciences[1]. Epidemiological applications have been anticipated for disease systems where the host is a plant or volume of soil[2,3], and hence is fixed in space. However, no natural examples have been reported. The central question of interest in percolation theory[4], the possibility of an infinite connected cluster, corresponds in infectious disease to a positive probability of an epidemic. Archived records of plague (infection with *Yersinia pestis*) in populations of great gerbils (*Rhombomys opimus*) in Kazakhstan have been used to show that epizootics only occur when more than about 0.33 of the burrow systems built by the host are occupied by family groups[5]. The underlying mechanism for this abundance threshold is unknown. Here we present evidence that it is a percolation threshold, which arises from the difference in scale between the movements that transport infectious fleas between family groups and the vast size of contiguous landscapes colonized by gerbils. Conventional theory predicts that abundance thresholds for the spread of infectious disease arise when transmission between hosts is density dependent such that the basic reproduction number (R_0) increases with abundance, attaining 1 at the threshold. Percolation thresholds, however, are separate, spatially explicit thresholds that indicate long-range connectivity in a system and do not coincide with $R_0 = 1$. Abundance thresholds are the theoretical basis for attempts to manage infectious disease by reducing the abundance of susceptibles, including vaccination and the culling of wildlife[6–8]. This first natural example of a percolation threshold in a disease system invites a re-appraisal of other invasion thresholds, such as those for epidemic viral infections in African lions (*Panthera leo*), and of other disease systems such as bovine tuberculosis (caused by *Mycobacterium bovis*) in badgers (*Meles meles*).

In infectious disease epidemiology, there is an emphasis on thresholds that can be traced back to the threshold property of the basic reproduction number, R_0, defined here as the expected number of new infections arising from a single infectious host in a population of susceptibles. Values less than 1 imply that infectious individuals fail to replace themselves and that the disease inevitably dies out, whereas values greater than 1 imply that invasion of the disease into the host population is possible[9]. The difficulties with using reproduction numbers such as R_0 when a host population is spatially or socially structured are generally understood[10,11]. Nevertheless, they remain the theoretical context for interpreting empirically observed thresholds[12] including the abundance threshold for plague epizootics[5,13]. Here, though, we present evidence that the plague threshold is a natural percolation phenomenon, and cannot be interpreted as empirical support for conventional thresholds based on R_0.

To explain this, we begin with a sketch of percolation theory. We then describe aspects of the population biology of great gerbils, and the landscapes they inhabit, which suggest percolation theory is an appropriate approach to understanding plague epizootics. A network model is then elaborated. The purpose of this model is to bring together three spatial scales: (i) that of flea movements responsible for plague transmission between family groups of great gerbils; (ii) the dimensions of the contiguous landscapes inhabited by great gerbils; and (iii) the scale at which plague monitoring is conducted in Kazakhstan. We argue that the network model shows how the three spatial scales are together responsible for the empirical observation of an abrupt threshold for plague.

Percolation theory concerns the behaviour of connected clusters in random networks[4]. Whether an infectious disease will spread among a population of hosts that have a fixed position in space and may only infect their nearest neighbours is well recognized as a percolation problem[14,15]. Equally recognized is the relevance of percolation theory to epidemics on networks in general, and lattice models in particular[16–18]. Among empiricists, the theory has been used to postulate the existence of spatial thresholds for the spread of fungal parasites of plants[2].

A simple case of percolation is when bonds form independently between adjacent points on the plane square lattice with probability p. In network terminology, each point on the lattice is a vertex and if a bond is present there is said to be an open edge between the two vertices. A 'cluster' is then defined as a set of vertices connected by open edges. In percolation theory the questions of interest are whether an infinite cluster exists, whether this cluster is unique and whether a given vertex will belong to such a cluster (the existence of an infinite cluster does not imply that all vertices belong to it). The pertinent result is that there is a critical value of the probability p, denoted by p_c, below which every vertex belongs to a small (finite) cluster. For $p > p_c$, an infinite cluster exists and there is a (strictly) positive probability that a given vertex belongs to it. In the case of bond percolation on the plane square lattice, the infinite cluster is unique and p_c is equal to $\frac{1}{2}$ (ref. 19). For numerous other networks, such a p_c also exists, though its value depends on the geometry of the network. In epidemiological applications, the given vertex corresponds to a primary case of an infectious disease and then p_c is an invasion threshold because the size of the cluster that the primary case belongs to corresponds to the final size of the epidemic. That is, when $p < p_c$ the spread from the primary case is limited to a small cluster of hosts.

Great gerbils are desert rodents inhabiting vast areas of central Asia. They form family groups and rely on underground burrow systems to help protect them from predators and the extreme temperatures. The burrow systems vary in size but are often large, complex constructions

[1]Theoretical Epidemiology, Faculty of Veterinary Medicine, University of Utrecht, Yalelaan 7, 3584 CL Utrecht, The Netherlands. [2]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, PO Box 85500, 3508 GA Utrecht, The Netherlands. [3]Department of Biology, University of Antwerp, Groenenborgerlaan 171, B-2020 Antwerp, Belgium. [4]Danish Pest Infestation Laboratory, University of Aarhus, Faculty of Agricultural Sciences, Department of Integrated Pest Management, Skovbrynet 14, DK-2800 Kongens Lyngby, Denmark. [5]Host-Parasite Biology Research Group, School of Biological Sciences, University of Liverpool, Crown Street, Liverpool L69 7ZB, UK.

representing the efforts of many generations. The vegetation above and around the burrow systems disappears such that a disc of bare earth forms, typically 30 m in diameter. These discs, and the patterned landscapes they create, are highly visible on satellite images (Fig. 1). Such images bring into clarity the regular spatial arrangement of burrow systems and the extent of the landscapes inhabited by great gerbils. The abundance of great gerbils, expressed as variation in the proportion of burrow systems inhabited, fluctuates over time[20]. More often than not, less than half of the burrow systems are occupied by family groups.

Plague bacteria are transmitted between great gerbils by fleas (mostly of the genus *Xenopsylla*) that inhabit the burrow systems. So an infectious flea has easy access to great gerbils living in the same burrow system but relatively limited access to those in surrounding burrow systems. Classifying whole family groups as either susceptible, infectious or recovered has been shown empirically to be a meaningful way of describing the distribution of plague in a great gerbil population[21]. Hence, in the network model we now present, vertices represent family groups rather than individual great gerbils. The random networks are generated from a landscape of burrow systems having the same structural characteristics as those observed on satellite images: 2.05 burrow systems per hectare with a buffer of 30 m (see Supplementary Information). Networks are formed from this landscape by choosing (random) subsets of the burrow systems to be vertices occupied by family groups. The size of the random subset is determined by the abundance of the host population, measured as occupancy.

Transmission of plague between family groups relies on the transport of infectious fleas between burrow systems. It occurs when an infectious flea migrates to its burrow system entrance and successfully jumps to a passing gerbil. In the network model, then, a transmission event is when an infectious flea attempts to leave its burrow system. We assume a flea cannot differentiate between a gerbil that belongs to the burrow system it is trying to leave and a gerbil from another burrow system, so there is a chance the flea fails to disperse. Transmission events only occur at infected vertices and hence occur at a rate proportional to the number of infected vertices. If the vertex that receives the infectious flea is susceptible then it becomes infected. To determine the vertex that received the infectious flea, we used gerbil movement data recorded during mark–recapture studies[22], combined with older field studies of flea dispersal in which fleas were marked using radionucleotides and their movements observed directly[23,24], to construct a set of weights for the vertices surrounding the infected vertex, including the infected vertex itself (see Supplementary Information). Most (more than 95%) observed flea movements, and great gerbil movements, were less than 200 m.

Each sample of great gerbils, tested for plague, is linked to a 'sector' (10 km × 10 km area) in the archives. The PreBalkhash focus consists of approximately 352 such sectors, only a fraction of which are visited each surveillance season. The results of the network model at this scale are shown in Fig. 2. Of principal interest is how the probability of an epizootic varies with the abundance of great gerbils expressed as the proportion of burrow systems occupied (varied between 0.01 and 0.49). In the model, there are three possible outcomes when a single family group is infected: (i) the infection dies out within this first family group, failing to spread to any other; (ii) infection spreads but only within a small cluster of family groups (there is a minor outbreak); or (iii) infection spreads within a large cluster (epizootic). A percolation threshold emerges provided the criterion defining a large cluster is that plague spreads more than several kilometres from the

Figure 1 | The regular, star-like pattern created by burrow systems, visible on satellite images. Patches of bare earth form above and around the burrow systems dug by great gerbils and strongly reflect the sunlight. Each bright disc represents a burrow system 10–40 m in diameter. The image was captured using the publicly available software Google Earth (http://earth.google.com/). Copyright 2008 DigitalGlobe; Europa Technologies.

site of first infection. As the distance used to define a large cluster is increased, the curves in Fig. 2 approach a non-differentiable curve similar to that found for percolation on infinite lattices[4]; that is, there is an interval of p over which the chance of a large-scale plague epizootic increases abruptly as p increases. When epizootics are defined as spread of plague over shorter distances, the chance of an epizootic increases steadily rather than abruptly. The spatial scale at which an epizootic is defined to have occurred therefore plays a large role in whether results indicate an abrupt threshold.

The percolation threshold in the network model occurs at about 0.31. This may be compared to the abundance threshold directly estimated from the plague archives by noting that the estimated threshold in a single-abundance model[5], which performed almost as well as the best, dual-abundance model (a difference in Akaike's information criterion of just 0.71), was 0.33 (95% confidence interval: 0.287, 0.373). That is, the percolation threshold falls within the 95% confidence interval of the threshold estimated from the plague archives. Hence we may, notwithstanding that some parameters in the network model were set to orders of magnitude, at least conclude that there is certainly no conflict between the quantitative results
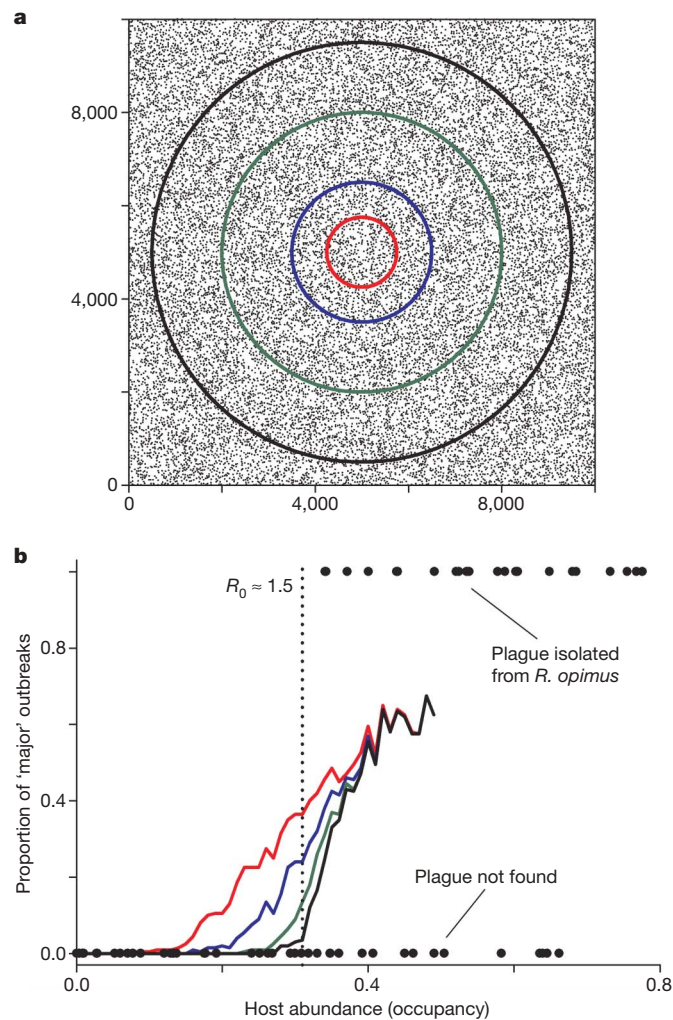
from the network model and the empirical abundance threshold, previously estimated. The sensitivity of the model to its parameters, including changes to the frequency of long-range flea movements, reveals a robust percolation result (see Supplementary Information).

A percolation threshold in an infectious-disease system takes into account the local depletion of susceptibles that arises from spatial restrictions on who can (or is likely to) infect whom. It is the point at which the landscape is sufficiently filled with hosts or host groups for the disease to continue to spread despite this barrier. This is distinctly different from the conventional understanding of invasion thresholds based on R_0 and models of randomly mixing populations. We measured R_0 directly from the network model simulations by recording the number of vertices infected by the first infected vertex and then averaging over many simulations (consistent with the definition of R_0 as the expected number of secondary cases resulting from a single primary case in a fully susceptible population). At an occupancy of 0.31, the value of occupancy at which the percolation threshold occurs, R_0 is around 1.5. More generally, for epidemics on lattice-like networks, the spatial constraints on who can infect whom mean that the expected number of secondary cases can be well above 1 but still only small outbreaks are possible[25]. How much larger than 1 R_0 must be for sustained spread to occur will depend on the details of particular systems, including the geometry of the spatial arrangement of hosts.

Our conclusion is that epizootics of sylvatic plague are a natural percolation phenomenon and that this is the mechanism (together with the spatial scale of surveillance in Kazakhstan) responsible for the abrupt population threshold seen in the plague archives. The empirical basis for this statement is the contrast between the lengths of movements of fleas between burrow systems and the scale of contiguous landscapes inhabited by the host. As the first natural example of a percolation threshold for disease invasion to be reported, it is even more interesting because the hosts are vertebrates. One consequence is a re-appraisal of other invasion thresholds such as the minimum densities of susceptible African lions for epidemic viral infections in the Serengeti[26]. The immune (or nearly immune) prides may create a network of hosts that is difficult for the aetiological agent to percolate through. This depends critically on whether transmission of the viruses between prides of lions occurs by (aggressive) contacts between lions from different prides, or whether mobile reservoir hosts such as spotted hyenas or nomadic lions are acting as a vector[27]. A second system that may warrant re-evaluation is bovine tuberculosis in badger populations, wherein badger setts may be analogous to great gerbil burrow systems. If percolation is the pertinent perspective, such that spread of tuberculosis in a badger population is dependent on long-range connectivity, then any lines of reasoning about the ability of badgers to sustain the disease (as a reservoir independent of transmission back and forth within the cattle population) must also be seen in that light.

Figure 2 | The results of the network model for plague epizootics in great gerbils. The landscape (**a**) is a 10 km × 10 km area, and the results (**b**) are expressed as the fraction of outbreaks that give rise to new infections at least 750 m, 1.5 km, 3 km and 4.5 km from the site of initial infection, shown as solid lines coloured red, blue, green and black, respectively, and corresponding to the spread of plague beyond the circles of the same colour in **a**. At the estimated percolation threshold, R_0 does not coincide with 1. Empirical observations (filled circles) of epizootics show an abrupt threshold at about 0.33 (ref. 5).

METHODS SUMMARY

Random networks of vertices representing occupied burrow systems were generated from a single landscape of burrow systems, the construction of which was based on the spatial arrangement of burrow systems observed on satellite images. Epizootics, or failed epizootics, were initiated on these networks by choosing the burrow system closest to the centre of the landscape to be infected. The spread of plague from this burrow system (vertex) was simulated by using a state vector and an adjacency matrix. The state vector had length equal to the number of occupied burrow systems (vertices), and each entry was either 0, 1 or 2, these values respectively indicating a susceptible, infected or recovered family group. Epizootics on the random networks essentially consisted of events occurring at infected vertices (the dispersal of an infected flea or the recovery of a family group from plague).

Time was modelled explicitly by defining the gap between events as a random variable with an exponential distribution and a rate proportional to the number of infected vertices. The consequences of events were dealt with as they occurred: that is, for each event, a random infected vertex was chosen at which the event happened and if a flea dispersal event occurred then which burrow system

received the flea was also decided. If the burrow system that received the flea was susceptible then it immediately became infected (that is, the entry in the state vector changed from 0 to 1). If the event was a recovery event then the entry in the state vector changed from 1 to 2. A simulation ended either when there were no 1 entries in the state vector or if the distance between a newly infected vertex and the initially infected vertex exceeded 4.5 km.

1. Sahimi, M. *Applications of Percolation Theory* (Taylor & Francis, London, 1994).
2. Otten, W., Bailey, D. J. & Gilligan, C. A. Empirical evidence of spatial thresholds to control invasion of fungal parasites and saprotrophs. *New Phytol.* **163**, 125–132 (2004).
3. Bailey, D. J., Otten, W. & Gilligan, C. A. Saprotrophic invasion by the soil-borne fungal plant pathogen *Rhizoctonia solani* and percolation thresholds. *New Phytol.* **146**, 535–544 (2000).
4. Grimmett, G. *Percolation* (Springer-Verlag, Berlin, 1999).
5. Davis, S. *et al.* Predictive thresholds for plague in Kazakhstan. *Science* **304**, 736–738 (2004).
6. Artois, M., Delahay, R., Guberti, V. & Cheeseman, C. Control of infectious diseases of wildlife in Europe. *Vet. J.* **162**, 141–152 (2001).
7. Caley, P., Hickling, G. J., Cowan, P. E. & Pfeiffer, D. U. Effects of sustained control of brushtail possums on levels of *Mycobacterium bovis* infection in cattle and brushtail possum populations from Hohotaka, New Zealand. *NZ Vet. J.* **47**, 133–142 (1999).
8. Barlow, N. D. The ecology of wildlife disease control: simple models revisited. *J. Appl. Ecol.* **33**, 303–314 (1996).
9. Diekmann, O. & Heesterbeek, J. A. P. *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis, and Interpretation* (John Wiley, Chichester, UK, 2000).
10. Keeling, M. J. The effects of local spatial structure on epdemiological invasions. *Proc. R. Soc. Lond. B* **266**, 859–867 (1999).
11. Trapman, P. *On Stochastic Models for the Spread of Infections.* PhD thesis, Vrije Univ. Amsterdam (2006).
12. Lloyd-Smith, J. O. *et al.* Should we expect population thresholds for wildlife disease? *Trends Ecol. Evol.* **20**, 511–519 (2005).
13. Davis, S. *et al.* Empirical assessment of a threshold model for sylvatic plague. *J. R. Soc. Interface* **4**, 649–657 (2007).
14. Grassberger, P. On the critical behaviour of the general epidemic process and dynamical percolation. *Math. Biosci.* **63**, 157–172 (1983).
15. Mollison, D. Spatial contact models for ecological and epidemic spread. *J. R. Statist. Soc. B* **39**, 283–326 (1977).
16. Keeling, M. J. & Eames, T. D. Networks and epidemic models. *J. R. Soc. Interface* **2**, 295–307 (2005).
17. Newman, M. E. J. The structure and function of complex networks. *Siam Rev.* **45**, 167–256 (2003).
18. Trapman, P. On analytical approaches to epidemics on networks. *Theor. Popul. Biol.* **71**, 160–173 (2007).
19. Kesten, H. The critical probability of bond percolation on the square lattice equals 1/2. *Commun. Math. Phys.* **74**, 41–59 (1980).
20. Kausrud, K. L. *et al.* Climatically driven synchrony of gerbil populations allows large-scale plague outbreaks. *Proc. R. Soc. B* **274**, 1963–1969 (2007).
21. Davis, S. *et al.* Plague metapopulation dynamics in a natural reservoir: the burrow-system as the unit of study. *Epidemiol. Infect.* **135**, 740–748 (2007).
22. Begon, M. *et al.* Epizootiological parameters for plague (*Yersinia pestis* infection) in a natural reservoir in Kazakhstan. *Emerg. Infect. Dis.* **12**, 268–273 (2006).
23. Rudenchik, Yu. V., Soldatkin, I. S., Severova, E. A., Mokiyevich, N. A. & Klimova, Z. I. Quantitative evaluation of the possibility of the territorial advance of epizooty of plague in the population of the *Rhombymus opimus* (northern Karakum). *Zool. Zh.* **46**, 117–123 (1967).
24. Korneyev, G. A. *The Quantitative Characteristics of Parasitic Exchange in Some Mammal Species Resulting from Simulation of Epizootics in Desert Biocenoses* (Microbe Institute, Saratov, 1968).
25. Mollison, D. The dependence of epidemic and population velocities on basic parameters. *Math. Biosci.* **107**, 255–287 (1991).
26. Packer, C. *et al.* Viruses of the Serengeti: patterns of infection and mortality in African lions. *J. Anim. Ecol.* **68**, 1161–1178 (1999).
27. Roelke-Parker, M. *et al.* A canine distemper virus epidemic in Serengeti lions (*Panthera leo*). *Nature* **379**, 441–445 (1996).

Author Contributions S.D. proposed and pursued the question of what mechanism generated the plague threshold, constructed and analysed the network model and wrote the paper. P.T. introduced S.D. to percolation theory and contributed to the writing of the Supplementary Information. All authors discussed percolation and the plague system and made comments on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to S.D. (S.A.Davis@uu.nl).