LETTERS

Adaptation and the evolution of parasite virulence in a connected world

Geoff Wild[1], Andy Gardner[2] & Stuart A. West[3]

Adaptation is conventionally regarded as occurring at the level of the individual organism, where it functions to maximize the individual's inclusive fitness[1–3]. However, it has recently been argued that empirical studies on the evolution of parasite virulence in spatial populations show otherwise[4–7]. In particular, it has been claimed that the evolution of lower virulence in response to limited parasite dispersal[8,9] provides proof of Wynne-Edwards's[10] idea of adaptation at the group level. Although previous theoretical work has shown that limited dispersal can favour lower virulence, it has not clarified why, with five different suggestions having been given[6,8,11–15]. Here we show that the effect of dispersal on parasite virulence can be understood entirely within the framework of inclusive fitness theory. Limited parasite dispersal favours lower parasite growth rates and, hence, reduced virulence because it (1) decreases the direct benefit of producing offspring (dispersers are worth more than non-dispersers, because they can go to patches with no or fewer parasites), and (2) increases the competition for hosts experienced by both the focal individual ('self-shading') and their relatives ('kin shading'). This demonstrates that reduced virulence can be understood as an individual-level adaptation by the parasite to maximize its inclusive fitness, and clarifies the links with virulence theory more generally[16].

Darwin's theory of evolution by natural selection explains both the process and the purpose of adaptation[17,18]. The process of adaptation occurs through the action of natural selection, which is mediated by differential reproductive success of individual organisms, and resulting changes in gene frequency[17]. This process leads individual organisms to appear designed as if for the purpose of maximizing their inclusive fitness, which is defined as the effect of one individual's actions on its genetic contribution to future generations through its direct descendants and those of its relatives[1,2]. The inclusive fitness approach to adaptation has been extremely successful, especially in the fields of behavioural and evolutionary ecology, providing explanations for a wide range of traits[19,20].

Despite the success of inclusive fitness theory, a number of recent papers have challenged the idea, arguing that natural selection can favour group adaptations in cases in which inclusive fitness is not maximized[4–7]. This suggestion is analogous to Wynne-Edwards's original idea of group selection[10], whereby adaptations occur for the benefit of the group. The primary empirical evidence upon which this challenge is based[4–7] is the experimental observation that parasites (viruses) of both moths and bacteria evolve to cause less damage to their hosts (lower virulence) in spatially structured populations, where dispersal can be limited[8,9]. The argument here is that the parasites become more prudent to prevent overexploitation and, hence, avoid causing the extinction of the local host population. However, it seems plausible that this effect of limited dispersal could also be explained by inclusive fitness theory, because it will lead to a higher relatedness between interacting parasites, which has long been known to favour a more prudent exploitation of host resources and, therefore, a lower virulence[16] (Supplementary Information). The only way to resolve this debate is to move away from verbal arguments and towards formal theoretical models that incorporate explicit spatial dynamics such as variable patch sizes and within-patch demography, and to use such models to determine the underlying evolutionary mechanisms[21].

Here we address this problem by using a standard epidemiological model[16,22], in the context of a geographically structured population, to determine why limited parasite dispersal selects for lower levels of virulence. We assume the simplest possible situation to make the underlying selective forces explicit and to allow comparison with previous models, which have shown that dispersal influences virulence but have failed to clarify why[11–15]. More general discussion of the various ways in which virulence theory has been expanded, to examine the consequences of a range of potentially important biological factors, are provided elsewhere[16,23]. In addition to its role in the debate over the process of adaptation, this effect of dispersal may be particularly important for the evolution of parasites, because it suggests that as human activity makes the world more connected, natural selection will favour more virulent and dangerous parasites[12].

We assume an 'island model' with an infinite number of patches (subpopulations), each of which may contain up to N host individuals. In this model, an individual (host or parasite) either remains on its natal patch or disperses. If it disperses, each of the other patches in the population is an equally likely destination. The island model is a standard tool for examining the effect of population structure while allowing analytical simplifications by dividing interactions into those that are 'local' (same patch) and those that are 'global' (different patch)[21,24]. We assume that hosts reproduce at a constant per-capita rate, b. A newborn host will attempt to settle either on its natal patch (local dispersal), with probability $1-d_h$, or on a randomly chosen, non-natal patch (global dispersal), with probability d_h. A newborn host successfully settles on a patch only when the patch in question supports fewer than N individuals. If successful, the newborn host is assumed, for convenience, to mature instantaneously. If unsuccessful, the newborn dies. We assume also that adult hosts are not capable of dispersal—each adult remains on the patch it settled as a newborn.

We classify hosts as either infected by the parasite or uninfected. We ignore the possibility of multiple infections[16], so infected and uninfected hosts might also be called non-susceptible and susceptible, respectively. In our model, uninfected hosts die at constant per-capita rate, μ. Infected hosts, on the other hand, suffer a greater risk of mortality, dying at rate $\mu + z$. Here z describes the disease-induced mortality (parasite virulence) that arises as a consequence of the parasite's exploitation of its host.

We assume that parasite transmissibility, $\beta(z)$, is positively correlated with parasite virulence (z), to reflect the standard assumption

[1]Department of Applied Mathematics, University of Western Ontario, London, Ontario N6A 5B7, Canada. [2]Institute of Evolutionary Biology, University of Edinburgh, King's Buildings, Edinburgh EH9 3JT, UK. [3]Department of Zoology, University of Oxford, Oxford OX1 3PS, UK.

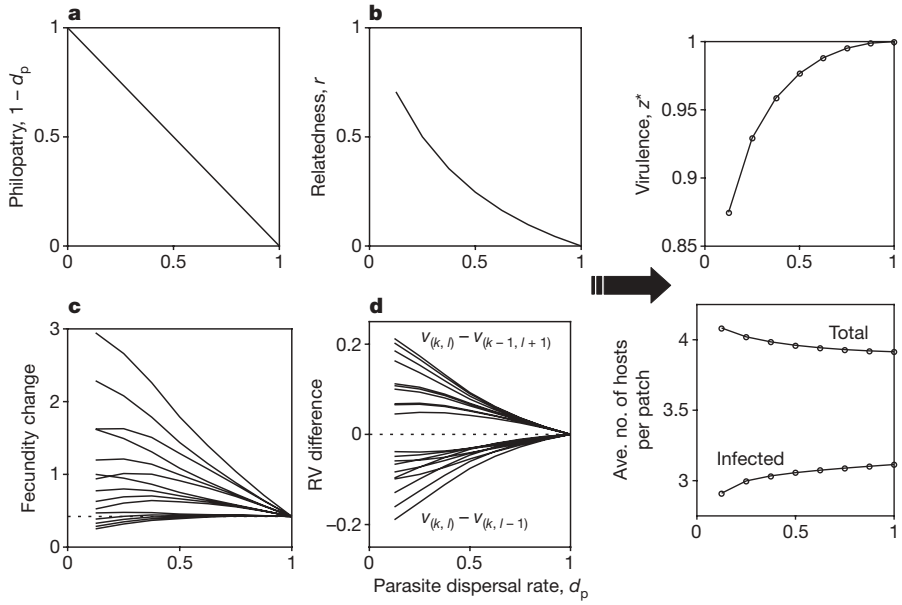

Figure 1 | Increasing parasite dispersal affects equation (1) in different ways. Increasing d_p decreases parasite philopatry (**a**); decreases the relatedness between parasite neighbours ($r_{(i,j)} = r$ for all i and $j > 2$; see Supplementary Information) (**b**); reduces the variation in parasite fecundity change across patch types (**c**); and reduces variation in reproductive value (RV) across parasite classes (**d**). The net effect of changes **a**–**d** is illustrated on the far right. Increasing d_p increases both the evolutionarily stable virulence level, z^*, and the fraction of hosts infected by the parasite. Results were generated by numerical simulation with $\mu = 1$, $d_h = 0.9$, $b = 3$, $N = 5$ and $\beta(x) = 5x/(1 + x)$.

that increased parasite growth leads not only to greater transmission, but also to greater host mortality[16,22]. We allow only horizontal transmission of the parasite, from infected adult to uninfected adult; hence, vertical transmission from host parent to its unborn offspring is not possible. Transmission is assumed to occur locally (within each patch), at a rate proportional to $(1-d_p)\beta(z)$, and globally (to other patches), at a rate proportional to $d_p\beta(z)$. In both cases, parasite transmission is assumed to follow a law of mass action. The parameter d_p is a proportion and is interpreted as the rate at which parasite offspring 'disperse' to new, randomly chosen patches.

We classify patches, and the parasites on those patches, according to the local number of uninfected (i) and infected (j) hosts. Naturally, parasite fitness depends upon the class to which its patch belongs, and the distribution of the different classes of patch in the population. To determine the evolutionarily stable level of parasite virulence, z^*, we consider a rare mutant parasite (the focal individual) belonging to class (k, l)—that is, on a patch with k uninfected hosts and l infected hosts. We note that although the global frequency of the mutant parasite is negligible, the probability that a parasite neighbour of a mutant is itself a mutant is not necessarily negligible. Thus, it is reasonable to expect that the social effects of mutant virulence are felt by other mutants as well.

In Supplementary Information, we show that if the focal mutant parasite increases its virulence phenotype by a small amount, $\delta > 0$, the resulting change in its inclusive fitness, $\Delta W_{(k,l)}$, is given by

$$\begin{aligned}\Delta W_{(k,l)} = &-\delta v_{(k,l)}\\ &+\delta\beta'(z)\left[(1-d_p)kv_{(k-1,l+1)} + d_p\sum\nolimits_{(i,j)} v_{(i,j)}(i+1)p_{(i+1,j-1)}\right]\\ &-\delta\beta'(z)(1-d_p)k(v_{(k,l)} - v_{(k-1,l+1)})\\ &-\delta\beta'(z)(1-d_p)k(v_{(k,l)} - v_{(k-1,l+1)})r_{(k,l)}(l-1)\\ &+\delta(v_{(k,l-1)} - v_{(k,l)})r_{(k,l)}(l-1)\end{aligned} \quad (1)$$

where $p_{(i,j)}$ is the equilibrium frequency of class-(i, j) patches, $r_{(k,l)} = r$ (Supplementary Information) is the relatedness between two different parasites on the same class-(k, l) patch, $v_{(k,l)}$ is the reproductive value[17] of a class-(k, l) parasite (the long-term genetic contribution made by such a parasite) and a prime denotes differentiation. Put verbally, equation (1) shows that the inclusive fitness effects of increased virulence are the cost of killing one's host, the benefits of enhanced transmission, the costs of increased competition for self, the costs of increased competition for relatives and the benefit to relatives due to killing one's host. In Supplementary Information, we show how equation (1) is used to determine the evolutionarily stable level of virulence (z^*).

In clear contrast to recent claims[4–7], analysis of equation (1) shows that the effect of parasite dispersal on virulence can be explained entirely using inclusive fitness theory (Fig. 1). Equation (1) is divided into the direct (personal) fitness consequences of increased virulence (first, second and third lines) and the indirect consequences for relatives (fourth and fifth lines). The first and second lines reflect the assumed compromise between host survival and parasite transmission: the host exploited by the mutant parasite suffers increased mortality (captured by the first term, $-\delta v_{(k,l)}$), whereas the parasite is able to produce new infections—both locally and globally—at a slightly higher rate (captured by

$$\delta\beta'(z)\left[(1-d_p)kv_{(k-1,l+1)} + d_p\sum\nolimits_{(i,j)} v_{(i,j)}(i+1)p_{(i+1,j-1)}\right]$$

which is called the 'fecundity change' term). For the special case of a well-mixed parasite population ($d_p = 1$), all but the first line of equation (1) vanishes (Supplementary Information), giving us the standard result that virulence evolves to maximize the basic reproductive number of the parasite[16,22].

The third line of equation (1) describes the direct (personal) fitness consequences, for the mutant, of increased local competition for fewer uninfected hosts. The increased transmissibility of the mutant increases the rate at which uninfected hosts become infected on the mutant's patch: class-(k, l) mutants move to class $(k-1, l+1)$ at a higher rate. Numerical results indicate that $v_{(k,l)} - v_{(k-1,l+1)} > 0$, so this change in the local host population represents an additional direct fitness cost of increased virulence that occurs in structured populations (Fig. 1d). Put simply, if parasite offspring do not disperse, then they decrease the local availability of uninfected hosts and increase the number of parasites competing for them. Consequently, increased parasite dispersal favours higher virulence because it reduces the direct

cost of producing offspring with which the focal individual will have to compete (a reduction in both the difference in reproductive values and in $1 - d_p$ in the third line of equation (1); Fig. 1a,d). This direct cost of parasite offspring production appears to be what has been described as self-shading[12]; thus, the third line of equation (1) can be thought of as a mathematical description of reduced self-shading due to parasite dispersal, but is also analogous to the 'tragedy of the commons'[8].

The fourth and fifth lines of equation (1) describe how the increased virulence exhibited by a mutant also has indirect fitness consequences, through changes to the competitive environment experienced by relatives. The major effect is that the increased transmission that results from the higher virulence of the mutant means that the relatives of the mutant also have increased competition for fewer locally available, uninfected hosts. Line four of equation (1) is simply the third line multiplied by both the number of parasites (other than the mutant actor) on the patch $(l-1)$ and the mean relatedness of those other parasites to the mutant ($r_{(k,l)}$). This indirect cost of increased virulence is reduced by parasite dispersal, through making relatives less likely to interact (Fig. 1b), and by decreasing the extent to which an increased virulence reduces the availability of uninfected hosts to relatives (a reduction in $1 - d_p$, $r_{(k,l)}$ and the difference in reproductive values in the fourth line of equation (1); Fig. 1a, b, d). Consequently, increased parasite dispersal favours higher virulence, because it reduces competition between relatives and, hence, reduces the indirect cost of higher virulence. This is analogous to self-shading but applies to the relatives of the mutant actor, so it could be thought of as kin shading. Kin shading is a between-hosts equivalent to the previous result that a lower relatedness (higher strain diversity) within the same host favours higher parasite virulence because it selects for faster growth rates, to obtain a higher proportion of the host resources[16].

The fifth line of equation (1) shows that the increased host mortality due to increased virulence affects the competitive environment experienced by the relatives of the mutant. Increased host mortality benefits relatives because it leads to a reduction in the number of locally competing parasites (a parasite dies along with its host) and because it clears a space that can later be filled by newborn (uninfected) hosts. Increased host mortality is also potentially costly to relatives, because it reduces the number of local hosts (a source of newborn, uninfected hosts). In many cases, numerical results indicate that $v_{(k,l-1)} - v_{(k,l)} > 0$, meaning that the increased host mortality that results from increased mutant virulence provides a net benefit to the mutant's relatives (upper lines in Fig. 1d). In these same cases, reduced parasite dispersal leads to a decrease in the competition experienced by relatives and, hence, an indirect benefit to higher virulence. In other cases, $v_{(k,l)} > v_{(k,l-1)}$ and the term $\delta(v_{(k,l-1)} - v_{(k,l)})r_{(k,l)}(l-1)$ of equation (1) counts as a cost in increased virulence. These latter cases are characterized by low host dispersal rates, so the cost of increased virulence here stems from the depletion of the main source of new, uninfected hosts. We must emphasize that even when $\delta(v_{(k,l-1)} - v_{(k,l)})r_{(k,l)}(l-1)$ counts as an inclusive fitness benefit, its size at equilibrium appears to be insufficient to raise the evolutionarily stable virulence level above that found in well-mixed populations.

More generally, as well as clarifying why the parasite dispersal rate should influence virulence, our model also shows how and why the parasite dispersal rate will interact with other parameters such as the maximum transmissibility of the parasite (β_{max}), the reproductive rate of the host and the host dispersal rate (Fig. 2). Increased host dispersal would favour increased virulence, through decreasing the extent to which increased virulence leads to self-shading and kin competition, as well as through any influence on parasite dispersal by moving parasites within hosts.

The reason why the parasite dispersal rate should influence virulence has proved controversial. Previous studies have offered four different explanations: virulent strains being surrounded by other infected individuals (self-shading)[12]; over-exploitation of the local availability of hosts (tragedy of the commons)[8]; competition between related strains ('kin selection')[15]; and the over-exploitation of local hosts and, hence, the extinction of parasite groups (the original Wynne-Edwards theory of 'group selection')[6,11]. It has also been suggested that the relationship between parasite dispersal and virulence is beyond the scope of existing evolutionary theory[25].

Our results show that an increase in parasite dispersal rate leads to selection for increased growth and, hence, to higher virulence for three reasons (Fig. 1). Increased dispersal provides a direct benefit to greater virulence, because it (1) increases the relative value of producing offspring (dispersers are worth more than non-dispersers) and (2) reduces the extent to which producing offspring will lead to the focal individual experiencing an increase in competition for available hosts (self-shading[12]). Increased dispersal provides an indirect benefit to greater virulence, because it (3) reduces the extent to which producing offspring will lead to relatives experiencing an increase in competition for available hosts (kin shading). The previous verbal explanations can be linked to these causal forces, in that self-shading[12] is our reason 1, the tragedy of the commons[8] involves our reasons 2 and 3, competition between relatives[15] is our reason 3 and the extinction of parasite groups[6,11] is linked to reasons 2 and 3; if an individual causes harm to their patch, then this cost is paid by both the focal individual and their relatives on the patch (that is, the group-selection components can always be partitioned into offspring and non-offspring components). There is also a fourth effect that works in the opposite direction, favouring a lower virulence with increased parasite dispersal. This effect is a consequence of the indirect benefit of reduced competition due to the number of relatives dying being greater at lower dispersal rates, but this is outweighed by the other three factors.

To conclude, we have shown that selection on rare mutant virulence phenotypes in structured populations of parasites can be explained by inclusive fitness theory. This is the latest of numerous

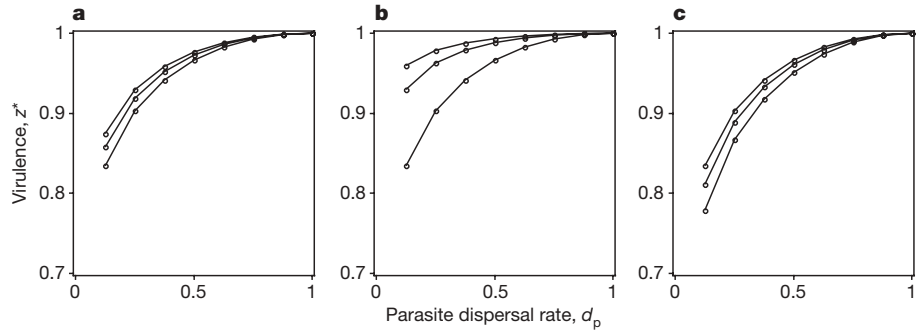

Figure 2 | Host and parasite life histories affect the relationship between stable virulence level and parasite dispersal rate. Relationship between z^* and d_p as host life-history parameters vary (**a**, **b**) and as maximum disease transmissibility, a parasite life-history trait, varies (**c**). **a**, From top to bottom, $d_h = 0.9$, 0.6 and 0.3 ($b = 3$). **b**, From top to bottom, $b = 9$, 6 and 3 ($d_h = 0.3$). Remaining parameters in **a** and **b** were $\mu = 1$, $N = 5$ and $\beta(x) = 5x/(1 + x)$. **c**, From top to bottom, $\beta_{max} = 5$, 7.5 and 20; remaining parameters were $\mu = 1$, $d_h = 0.3$, $b = 3$, $N = 5$ and $\beta(x) = \beta_{max}x/(1 + x)$. Additional, qualitatively similar results are presented in Supplementary Information.

examples that have accumulated over the past 30 years, in which it has been claimed that group selection and not kin selection is acting in a particular situation, only for explicit analyses to show otherwise[26,27]. Future confusion could be avoided if such claims were backed by formal analyses that actually examine the underlying selective forces, rather than just verbal arguments[26]. More generally, our results emphasize the difference between levels of adaptation and levels of selection[28]. The multilevel (group) selection and kin selection (inclusive fitness) approaches to social evolution have long been known to be mathematically equivalent and, if the analyses are performed correctly, do not lead to conflicting predictions[29,30]. Thus, irrespective of the relative strengths of within-group versus between-group selection, individuals are predicted to maximize their inclusive fitness. In contrast, groups are only predicted to evolve traits that function to maximize their fitness in extreme situations where there is no conflict of interest between the members of the group[28]. Put another way, the presence of group selection does not invalidate the idea that the individual is an adaptive unit, and it does not validate the idea that the group is an adaptive unit[28].

Received 13 January; accepted 9 April 2009.
Published online 27 May 2009.

1. Hamilton, W. D. The genetical evolution of social behaviour. *J. Theor. Biol.* **7,** 1–52 (1964).
2. Grafen, A. Optimization of inclusive fitness. *J. Theor. Biol.* **238,** 541–563 (2006).
3. Grafen, A. The formal Darwinism project: a mid-term report. *J. Evol. Biol.* **20,** 1243–1254 (2007).
4. Sober, E. & Wilson, D. S. *Unto Others: The Evolution and Psychology of Unselfish Behavior* (Harvard Univ. Press, 1998).
5. Wilson, D. S. & Wilson, E. O. Rethinking the theoretical foundation of sociobiology. *Q. Rev. Biol.* **82,** 327–348 (2007).
6. Wilson, D. S. Social semantics: towards a genuine pluralism in the study of social behaviour. *J. Evol. Biol.* **21,** 368–373 (2008).
7. Hölldobler, B. & Wilson, E. O. *The Superorganism: The Beauty, Elegance, and Strangeness of Insect Societies* (Norton, 2008).
8. Kerr, B., Neuhauser, C., Bohannan, J. M. & Dean, A. M. Local migration promotes competitive restraint in a host–pathogen 'tragedy of the commons'. *Nature* **442,** 75–78 (2006).
9. Boots, M. & Mealor, M. Local interactions select for lower pathogen infectivity. *Science* **315,** 1284–1286 (2007).
10. Wynne-Edwards, V. C. *Animal Dispersion in Relation to Social Behaviour* (Oliver & Boyd, 1962).
11. Haraguchi, Y. & Sasaki, A. Host-parasite arms race in mutation modifications: indefinite escalation despite a heavy load. *J. Theor. Biol.* **183,** 121–137 (1996).
12. Boots, M. & Sasaki, A. 'Small worlds' and the evolution of virulence: infection occurs locally and at a distance. *Proc. R. Soc. Lond. B* **266,** 1933–1938 (1999).
13. O'Keefe, K. J. & Antonovics, J. Playing by different rules: the evolution of virulence in sterilizing pathogens. *Am. Nat.* **159,** 597–605 (2002).
14. Boots, M., Hudson, P. J. & Sasaki, A. Large shifts in pathogen virulence relate to host population structure. *Science* **303,** 842–844 (2004).
15. Lion, S. & van Baalen, M. Self-structuring in spatial evolutionary ecology. *Ecol. Lett.* **11,** 277–295 (2008).
16. Frank, S. A. Models of parasite virulence. *Q. Rev. Biol.* **71,** 37–78 (1996).
17. Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon, 1930).
18. Hamilton, W. D. Selfish and spiteful behaviour in an evolutionary model. *Nature* **228,** 1218–1220 (1970).
19. Stearns, S. C. *Evolution of Life Histories* (Oxford Univ. Press, 1992).
20. Krebs, J. R. & Davies, N. B. *An Introduction to Behavioural Ecology* 3rd edn (Blackwell Scientific, 1993).
21. Rousset, F. & Ronce, O. Inclusive fitness for traits affecting metapopulation demography. *Theor. Popul. Biol.* **65,** 127–141 (2004).
22. Anderson, R. M. & May, R. M. Coevolution of hosts and parasites. *Parasitology* **85,** 411–426 (1982).
23. Day, T. & Gandon, S. Applying population-genetic models in theoretical evolutionary epidemiology. *Ecol. Lett.* **10,** 876–888 (2007).
24. Rousset, F. *Genetic Structure and Selection in Subdivided Populations* (Princeton Univ. Press, 2004).
25. Goodnight, C. *et al.* Evolution in spatial predator–prey models and the "prudent predator": the inadequacy of steady-state organism fitness and the concept of individual and group selection. *Complexity* **13,** 23–44 (2008).
26. Lehmann, L. & Keller, L. The evolution of cooperation and altruism – a general framework and a classification of models. *J. Evol. Biol.* **19,** 1365–1376 (2006).
27. West, S. A., Griffin, A. S. & Gardner, A. Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J. Evol. Biol.* **20,** 415–432 (2007).
28. Gardner, A. & Grafen, A. Capturing the superorganism: a formal theory of group adaptation. *J. Evol. Biol.* **22,** 659–671 (2009).
29. Frank, S. A. Hierarchical selection theory and sex ratios. I. General solutions for structured populations. *Theor. Popul. Biol.* **29,** 312–342 (1986).
30. Queller, D. C. Quantitative genetics, inclusive fitness, and group selection. *Am. Nat.* **139,** 540–558 (1992).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Boots, S. Brown, A. Buckling, A. Fernandes, S. Gandon, A. Griffin and S. Lion for discussions, and the Natural Sciences and Engineering Research Council of Canada, the UK Natural Environmental Research Council, the Royal Society, the Leverhulme Trust, and the European Research Council for funding.

Author Contributions All authors contributed equally to this work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to G.W. (gwild@uwo.ca).

Patch Dynamics

Habitat patches, in our model, are classified according the number of susceptible hosts (i) and the number of infected hosts (j) found to reside there (here, $0 \le i + j \le N$). We use the vector, (i, j) to indicate the class to which a given patch belongs.

Let $x_{(i,j)}$ denote the density of class-(i,j) patches in our infinite model population, and let $p_{(i,j)} = x_{(i,j)} \Big/ \sum_{(k,l)} x_{(k,l)}$ denote the population-wide frequency of class-(i, j) patches.

Class- (i,j) patches are created from other patches, whose classification belongs to the set

$$\mathbf{U}_{(i,j)} = \{(i-1,j),(i+1,j),(i,j+1),(i+1,j-1)\}$$

as hosts either die or become infected. Class-(i,j) patches are destroyed in the same way to become patches whose classification belongs to the set

$$\mathbf{D}_{(i,j)} = \{(i-1,j),(i+1,j),(i,j-1),(i-1,j+1)\}.$$

If $q_{(i,j),(k,l)}$ denotes the rate at which a class-(k,l) patch becomes a class-(i,j) patch, then the *patch dynamics*, are described by

$$\frac{dx_{(i,j)}}{dt} = \sum_{(k,l)\in\mathbf{U}_{(i,j)}} q_{(i,j),(k,l)} x_{(k,l)} - \sum_{(k,l)\in\mathbf{D}_{(i,j)}} q_{(k,l),(i,j)} x_{(i,j)}. \qquad \text{(S1)}$$

In Table S1, we state the rates $q_{(i,j),(k,l)}$, using the notation introduced in the main text of the paper, as well as the shorthand notation,

$$b_{(k,l)} = \begin{cases} b(1-d_h)(k+l) + bd_h(S+I) & \text{for } k+l < N \\ 0 & \text{otherwise} \end{cases}$$

where b is the per capita rate of host reproduction (for both susceptible and infected hosts), d_h is the natal dispersal rate of the host, and

$$S = \sum_{(k,l)} kp_{(k,l)} = \sum_{k=0}^{N}\sum_{l=0}^{N-k} kp_{(k,l)},$$

$$I = \sum_{(k,l)} lp_{(k,l)} = \sum_{k=0}^{N}\sum_{l=1}^{N-k} lp_{(k,l)} = \sum_{k=0}^{N}\sum_{l=0}^{N-k} lp_{(k,l)}.$$

Additional shorthand notation includes,

$$\beta_{(k,l)}(z) = \beta(z)\left[\left(1-d_p\right)l + d_p I\right]k,$$

and

$$\tilde{\beta}_{(k,l)}(z) = \begin{cases} \beta(z)d_p(k+1)p_{(k+1,l-1)} & k+l \le N, \\ 0 & otherwise. \end{cases}$$

Since $\sum_{(i,j)} dx_{(i,j)}/dt = 0$ (patches are neither created, nor destroyed) we can say that

$$\frac{dp_{(i,j)}}{dt} = \sum_{(k,l)\in \mathbf{U}_{(i,j)}} q_{(i,j),(k,l)}p_{(k,l)} - \sum_{(k,l)\in \mathbf{D}_{(i,j)}} q_{(k,l),(i,j)}p_{(i,j)}. \qquad \text{(S2)}$$

The inclusive fitness analysis we undertake in the next section (and, indeed, in the main text) assumes that the solution of equation (S2) tends toward a unique, globally stable equilibrium. From this point forward, then, $\mathbf{p} = (p_{(i,j)})_{(i,j)}$ will refer only to this equilbrium state. We suppose further that classes (i, j) such that $j \ge 1$ are in the support of $\mathbf{p}$. Unfortunately, non-trivial equilibrium solutions of equation (S1) can only be found by numerical methods. Nevertheless, numerical investigation indicates that our assumption, above, holds for a wide range of parameter values.

Inclusive Fitness Model

In this section, we give a more detailed description of how we reach equation (1) of the main text. The approach is the same as that used recently by Alizon & Taylor[31] (see also Appendix A of ref. 32). We begin by assuming a monomorphic parasite population, i.e. a parasite population in which all individuals employ the same virulence "strategy" z.

Both the class to which a parasite belongs, and the distribution of the different patch types in the population have important implications for parasite fitness. Let $w_{(i,j),(k,l)}$ denote the (i,j) fitness of class-(k,l) parasite, defined as the per capita rate at which class-(k,l) pathogens produce class-(i,j) infections, when the distribution of patch types has reached an equilibrium—one that includes patches with infected individuals. Expressions for $w_{(i,j),(k,l)}$ are summarized in Table S2. From these "fitness functions" we can determine $u_{(i,j)}$, the equilibrium frequency of class-(i,j) parasites, as well as $v_{(i,j)}$ the individual reproductive value of class-(i,j) parasites (i.e. the per capita contribution of class-(i,j) to the generations in the distant future). The former quantities satisfy the equation, $\sum_{(k,l)} w_{(i,j),(k,l)} u_{(k,l)} = 0$, whereas the latter quantities satisfy, $\sum_{(i,j)} v_{(i,j)} w_{(i,j),(k,l)} = 0$. Equilibrium frequencies and reproductive values, as we will see, are some of the key building blocks of kin selection models[32, 33].

We have, so far, been thinking of a scenario in which all parasites exploit hosts to the same extent. In this scenario, all infected hosts suffer "normal" disease-induced mortality at rate, z. To build our kin selection model, we introduce slightly deviant and extremely rare exploitation strategy that changes the mortality rate of infected hosts by a very small amount[34,35]. We then ask, does a shift from a normal host exploitation strategy to a deviant one increase or decrease the inclusive fitness of a parasite?

Let $\Delta W_{(k,l)}$ denote the change that occurs in the inclusive fitness of a single class-(k,l) pathogen (the "focal actor") as it changes its exploitation strategy from normal to deviant. As mentioned above, the strategy shift implies that the mortality of the individual playing host to the focal actor changes by a small amount, which we now denote as, $\delta > 0$. We record this increased mortality as a cost paid by the focal actor and each of its $(l-1)$ patchmates, writing $-\delta v_{(k,l)}\left(1+(l-1)r_{(k,l)}\right) = -\delta v_{(k,l)}\bar{r}_{(k,l)} l$, where $r_{(k,l)}$ is the relatedness between two different parasites on the same class-(k,l) patch, and $\bar{r}_{(k,l)}$ is the relatedness between two pathogens on the same class-(k,l) patch, chosen at random with replacement. The term $v_{(k,l)}$ reflects the fact that the costs are in (k,l)-fitness and must be included to make sure all fitness changes receive proper weighting in the final calculation.

Increased host mortality also carries with it certain inclusive fitness benefits when $l > 1$. When the focal actor and its host have died, $(l-1)$ surviving parasites immediately join class-$(k, l-1)$. We record this change as, $\delta v_{(k,l-1)} r_{(k,l)}(l-1) = \delta v_{(k,l-1)}\left(\bar{r}_{(k,l)} l - 1\right)$.

When $k \geq 1$ susceptible hosts can be found locally, and so increased host exploitation implies that the focal actor produces new local infections at a higher rate. In this case, the rate at which local infections occur is increased by $\delta\beta'(z)\left(1-d_p\right)k$. We record the inclusive-fitness change as $\delta v_{(k-1,l+1)}\beta'(z)\left(1-d_p\right)k\left(1+\bar{r}_{(k,l)}l\right)$. The multiplier, $\left(1+\bar{r}_{(k,l)}l\right)$ is included here, because, when a new local infection occurs, the actor, the actor's "offspring" and the actor's $(l-1)$ patchmates all become class-(k–1, l+1) parasites simultaneously. Naturally, when the actor and its patchmates become class-(k–1, l+1) parasites they cease to be class-(k,l) pathogens, hence we must record an additional inclusive fitness loss of, $-\delta v_{(k,l)}\beta'(z)\left(1-d_p\right)kl\bar{r}_{(k,l)}$ in our calculation.

Increased host exploitation also means that the rate at which the focal parasite produces non-local infections increases. Specifically, if $p_{(i+1,j-1)}$ denotes the frequency of class-$(i+1, j-1)$ patches, the increase in $w_{(i,j)}$ is given by, $\delta v_{(i,j)}\beta'(z)d_p\left(i+1\right)p_{(i+1,j-1)}$. We make the standard assumption that non-local infections occur only on very distant patches, so that the effects these infections have on the fitness of (similarly distant) relatives of the focal parasite need not be included in our calculation. We can then write the change in the inclusive fitness of a class-(k,l) parasite as

$$\Delta W_{(k,l)} = -\delta v_{(k,l)}\bar{r}_{(k,l)}l + \delta v_{(k,l-1)}\left(\bar{r}_{(k,l)}l - 1\right) + \delta v_{(k-1,l+1)}\beta'(z)\left(1-d_p\right)k\left(1+\bar{r}_{(k,l)}l\right)$$
$$-\delta v_{(k,l)}\beta'(z)\left(1-d_p\right)kl\bar{r}_{(k,l)} + \delta\sum_{(i,j)} v_{(i,j)}\beta'(z)d_p\left(i+1\right)p_{(i+1,j-1)} \qquad \text{(S3)}$$

which rearranges to give equation (1) of the main text. The overall change in the inclusive fitness of the mutant, call it ΔW, is a weighted average of all $\Delta W_{(k,l)}$, where class frequencies $u_{(k,l)}$ are used as weights.

Calculation of Relatedness Coefficients

We can calculate the coefficients of relatedness for a neutral population using the notion of identity by descent, or IBD. Two genes are said to be IBD provided they have descended from a common ancestor without intervening mutation. The coefficient of consanguinity (CC) between two (same-locus) genes is simply the probability that the alleles are IBD.

Let $r_{(i,j)}$ with $j \geq 2$ denote the CC between genes drawn from two different class-(i, j) parasites found on the same patch. To calculate $r_{(i,j)}$, observe that a class-(i, j) patch was a class-(k, l) patch dt time units ago with probability $\alpha_{(i,j),(k,l)}dt$, where $\alpha_{(i,j),(k,l)} = \frac{q_{(i,j),(k,l)}p_{(k,l)}}{p_{(i,j)}}$. We use $\alpha_{(i,j),(k,l)}$ to write a differential equation that describes how the CCs $r_{(i,j)}$ change over time. Assuming $j > 2$, then

$$\frac{dr_{(i,j)}}{dt} = \alpha_{(i,j),(i-1,j)}\left(r_{(i-1,j)} - r_{(i,j)}\right) + \alpha_{(i,j),(i+1,j)}\left(r_{(i+1,j)} - r_{(i,j)}\right) + \alpha_{(i,j),(i,j+1)}\left(r_{(i,j+1)} - r_{(i,j)}\right)$$

$$+\alpha_{(i,j),(i+1,j-1)}\frac{\left(1-d_p\right)(j-1)}{\left(1-d_p\right)(j-1)+d_pI}\left[\frac{2}{j(j-1)} + \left(1-\frac{2}{j(j-1)}\right)r_{(i+1,j-1)} - r_{(i,j)}\right]$$

$$+\alpha_{(i,j),(i+1,j-1)}\frac{d_pI}{\left(1-d_p\right)(j-1)+d_pI}\left[\left(1-\frac{2}{j}\right)r_{(i+1,j-1)} - r_{(i,j)}\right]. \qquad \text{(S4)}$$

We omit the term $\alpha_{(i,j),(i-1,j)}\left(r_{(i-1,j)} - r_{(i,j)}\right)$ in (S4) when $i = 0$, and we omit the terms $\alpha_{(i,j),(i+1,j)}\left(r_{(i+1,j)} - r_{(i,j)}\right)$ and $\alpha_{(i,j),(i,j+1)}\left(r_{(i,j+1)} - r_{(i,j)}\right)$ when $i + j = N$. When $j = 2$, there are no terms in eqn (A6) that involve the undefined coefficient $r_{(i+1,\,j-1)}$. In this case, (S4) becomes

$$\frac{dr_{(i,2)}}{dt} = \alpha_{(i,2),(i-1,2)}\left(r_{(i-1,2)} - r_{(i,2)}\right) + \alpha_{(i,2),(i+1,2)}\left(r_{(i+1,2)} - r_{(i,2)}\right) + \alpha_{(i,2),(i,3)}\left(r_{(i,3)} - r_{(i,2)}\right)$$

$$+\alpha_{(i,2),(i+1,1)}\left[\frac{\left(1-d_p\right)}{\left(1-d_p\right)+d_pI} - r_{(i,2)}\right].$$

Remarkably, at equilibrium $r_{(i,j)} = r$ for all (i, j), where

$$r = \frac{\left(1 - d_p\right)}{\left(1 - d_p\right) + d_p I} \leq 1. \tag{S5}$$

Since $I > 0$, equality in (A8) holds only when $d_p = 0$. Notice that, as the average number of infections per patch (I) increases, r decreases. Notice also that

$$\bar{r}_{(i,j)} = \frac{1}{j} + \frac{j-1}{j} r.$$

The fact that r does not depend on j is worth discussing, since the result may initially appear counterintuitive. The absolute rates of host mortality and local infection within each patch are, of course, in proportion to the number of infected hosts found locally. However, at equilibrium, the impact of each mortality or infection event upon local relatedness is in inverse proportion o the number of local infected hosts. Hence, although relatedness-altering events occur more frequently in patches with more infected hosts, the individual impact of each event is more diluted.

Well-Mixed Pathogen Populations

When pathogen dispersal, $d_p = 1$ we say that the pathogen population is well-mixed. In this case, the class structure of the pathogen population has no bearing on pathogen fitness and pathogens are not related to non-self patchmates. Formally, we can say that $v_{(k,l)} \equiv 1$, $\bar{r}_{(k,l)} = 1/l$, and so

$$\Delta W = -\delta + \delta\beta'(z)S,$$

where $S = \sum_{(i,j)} (i+1) p_{(i+1,j-1)} = \sum_{(i,j)} i p_{(i,j)}$ gives the average number of susceptible individuals per patch. At evolutionary equilibrium, then, our inclusive fitness model tells us that the marginal fitness benefit due to increased in disease transmission (i.e. the term, $\delta\beta'(z)S$) is exactly balanced by the marginal fitness cost of increased host mortality (i.e. the term, $-\delta$). This same result can be established with a standard game

theory model that measures fitness using pathogen lifetime reproductive success, $\beta(z)S/(\mu+z)$ (refs 16,36).

If we assume the simple virulence-transmissibility relationship, $\beta(z)=\beta_{\max}z/(1+z)$, then disease transmissibility cannot exceed the maximum level given by $\beta_{\max}$. With this assumption it is possible to show that evolution is at equilibrium with respect z when $z=z^*=\sqrt{\mu}$. In particular when $\mu=1$, $z^*=1$.

Numerical Investigations

We analyzed the inclusive fitness model from the main text to identify those disease-induced mortality rates associated with the convergence stable levels of host exploitation. Convergence stability is a notion borrowed from evolutionary game theory[35,37,38]. If z^* is convergence stable, then the distance between z^* and x, the phenotype displayed by successful mutant invaders of a population that is otherwise fixed at z, is always less than the distance between z^* and z. In the context of the present model the disease-induced mortality rate associated with the convergence stable host exploitation rate is an evolutionary equilibrium, z^* that satisfies the following pair of conditions: $\Delta W\big|_{z=z^*}>0$ when $\delta<0$ and $\Delta W\big|_{z=z^*}<0$ when $\delta>0$.

Results were generated through an iterative numerical procedure. For each parameter combination, we began by guessing a corresponding z^*, call it z_0. Given z_0, we determined the sign of the inclusive fitness effect, ΔW and refined our guess accordingly. Given the refined guess, z_1, the sign of ΔW was determined yet again, and further refinements were made. The process of calculating the sign of ΔW and making smaller and smaller refinements continued until successive refinements agreed to several decimal places. We implement our numerical procedure using the computer software package, Maple (version 11).

Predictions made by the inclusive fitness model proposed here agree with those made elsewhere. Most importantly, our model predicts that reduced mixing of the pathogen population (i.e. lower dispersal rate, d_p) promotes reduced host exploitation, and in turn reduced disease-induced host mortality (Figs. S1, S2).

Naturally, the relationship between pathogen mixing and disease-induced host mortality depends on the parameter values under consideration. If there exists a maximum possible transmissibility (β_{max}), then increasing this maximum encourages larger reductions in disease-induced host mortality (Fig. S1). Moreover, the rate at which such reductions in host mortality occur depend not only β_{max} but also on N (Fig. S1).

Host reproduction rate (b) and, to a lesser degree, host dispersal rate (d_h) also influence the evolution disease-induced host mortality. Specifically, increasing the value of b and/or d_h reduces the extent to which parasite dispersal influences disease-induced mortality (Fig. S2).

Related Issues

We conclude by clarifying a number of links to existing work. Hamilton[39] was the first to suggest that a lower relatedness between the parasites infecting a host would lead to selection for faster exploitation of host resources, and hence a higher virulence, This has since been demonstrated formally by a number of authors[16,33,40-45]. Our results show that an analogous prediction (albeit more complex) arises in spatially structured populations, where patterns of dispersal determine the relatedness of parasites competing for hosts within patches (see also ref. 33, p.163-168). Our method builds upon previous applications of inclusive fitness theory that have examined the evolution of cooperation[31] in spatial populations with explicit within and between patch demography. In particular, we follow previous authors by treating the deviant virulence strategy as rare and by neglecting its long-term effects on both patch distribution and

reproductive value (see also Appendix A of ref. 32). In this way our analysis focuses on the initial success of a mutant, rather than its probability of fixation. It would certainly be very useful to examine the consequences of demographic stochasticity (e.g. as in ref. 21) or the consequences putting alternative forms of parasite interaction, such as cooperation[46,47], or spiteful interference[48] into an explicit spatial setting (see also ref. 49). Finally, we draw an analogy to the sex ratio literature, where the benefit of determining the underlying selective forces have long been appreciated[29,50].

Additional References

31. Alizon, S. & Taylor, P. D. Empty sites can promote altruistic behaviour. *Evolution* **62**,1335–1344 (2008).

32. Taylor, P. D. & Frank S. A. How to make a kin selection model. *J. Theor. Biol.* **176**, 27–37 (1996).

33. Frank, S. A. *Foundations of Social Evolution.* (Princeton University Press, Princeton, NJ, 1998).

34. Grafen, A. Hamilton's rule OK. *Nature* **318**, 310–311 (1985).

35. Taylor, P. D. Evolutionary stability in one-parameter models under weak selection. *Theor. Popul. Biol.* **36**, 125–143 (1989).

36. Day, T. & Burns, J. G. A consideration of patterns of virulence arising from host-parasite coevolution. *Evolution*, **57**, 671–676 (2003).

37. Eshel, I. Evolutionary and continuous stability. *J. Theor. Biol.* **103**, 99–111 (1983).

38. Christiansen, F. B. On conditions for evolutionary stability for a continuously varying character. *Am. Nat.* **138**, 37–50 (1991).

39. Hamilton, W.D. Altruism and related phenomena, mainly in social insects. *Annu. Rev. Ecol. Syst.* **3**, 193-232 (1972).

40. Bremermann, H. J. & Pickering, J. A game-theoretical model of parasite virulence. *J. Theor. Biol.* **100**, 411-26 (1983).

41. Frank, S. A. Kin selection and virulence in the evolution of protocells and parasites. *Proc. Roy. Soc. Lond. B* **258**, 153-61 (1994).

42. Nowak, M. & May, R. M. Superinfection and the evolution of parasite virulence. *Proc. Roy. Soc. Lond. B* **255**, 81-89 (1994).

43. May, R. M. & Nowak, M. A. Coinfection and the evolution of parasite virulence." *Proc. R. Soc. Lond. B* **261**, 209-15 (1995).

44. van Baalen, M. & Sabelis M. W. The dynamics of multiple infection and the evolution of virulence. *Am. Nat.* **146**, 881-910 (1995).

45. Gandon, S. & Michalakis, Y. Evolution of parasite virulence against qualitative or quantitative host resistance. *Proc. Roy. Soc. Lond. B* **267**, 985-990 (2000).

46. Brown, S. P., Hochberg, M. E. & Grenfell, B. T. Does multiple infection select for raised virulence? *Trends in Microbiology* **10**, 401-405 (2002).

47. West, S. A. & Buckling, A. Cooperation, virulence and siderophore production in bacterial parasites. *Proc. Roy. Soc. Lond. B* **270,** pp. 37-44 (2003).

48. Gardner, A.,West, S. A. & Buckling, A. Bacteriocins, spite and virulence." *Proc. Roy. Soc. Lond. B* **271**, 1529-2535 (2004).

49. Lion, S. & Gandon, S. Habitat saturation and the spatial evolutionary ecology of altruism. *J. Evol. Biol.*, in press.

50. Taylor, P D. Intra-sex and inter-sex sibling interactions as sex determinants. *Nature* **291**, 64-66 (1981).

Table S1. Rates $q_{(i,j),(k,l)}$ (from eqn S1) at which class-(i,j) patches are (a) created and (b) destroyed. Recall that i and j denote the number of susceptible and infected hosts, respectively, found on a given patch. We assume $i, j \geq 0$, with $i + j \leq N$

$(k,l) \rightarrow (i,j)$	Description	$q_{(i,j),(k,l)}$
(a) Inputs to class-(i, j)		
$(i+1, j) \rightarrow (i,j)$	susceptible host dies on (i+1,j)-patch	$\mu(i+1)$ (or zero if $i + j = N$)
$(i-1,j) \rightarrow (i,j)$	newborn host arrives on (i – 1,j)-patch	$b_{(i-1,\,j)}$
$(i, j+1) \rightarrow (i,j)$	infected host dies on (i, j+1)-patch	$(\mu + z)(j + 1)$ (or zero if $i+ j =N$)
$(i+1,j-1) \rightarrow (i,j)$	infection occurs on (i+1,j –1)-patch	$\beta_{(i+1,j-1)}(z)$
(b) Outputs from class-(i,j)		
$(i, j) \rightarrow (i-1, j)$	susceptible host dies on (i,j)-patch	μi
$(i, j) \rightarrow (i + 1, j)$	newborn host arrives on (i,j)-patch	$b_{(i,j)}$
$(i, j) \rightarrow (i, j-1)$	infected host dies on (i, j)-patch	$(\mu + z)j$
$(i,j) \rightarrow (i-1,j+1)$	infection occurs on (i,j)-patch	$\beta_{(i,j)}(z)$

Table S2. Expressions for wild-type, or "normal" fitness function, $w_{(i,j),(k,l)}$. Recall that $j, l \geq 1$ when describing pathogen class structure.

(i,j)	$w_{(i,j),(k,l)}$
(k, l)	$-\left(\mu k + (\mu + z)l + b_{(k,l)} + \beta_{(k,l)}(z)\right) + \tilde{\beta}_{(k,l)}(z)$
$(k-1, l)$	$\mu k + \tilde{\beta}_{(k-1,l)}(z)$
$(k+1, l)$	$b_{(k,l)} + \tilde{\beta}_{(k+1,l)}(z)$
$(k, l-1)$	$(\mu + z)(l-1) + \tilde{\beta}_{(k,l-1)}(z)$
$(k-1, l+1)$	$\beta(z)(1-d_p)k + \beta_{(k,l)}(z) + \tilde{\beta}_{(k-1,l+1)}(z)$
all others	$\tilde{\beta}_{(i,j)}(z)$

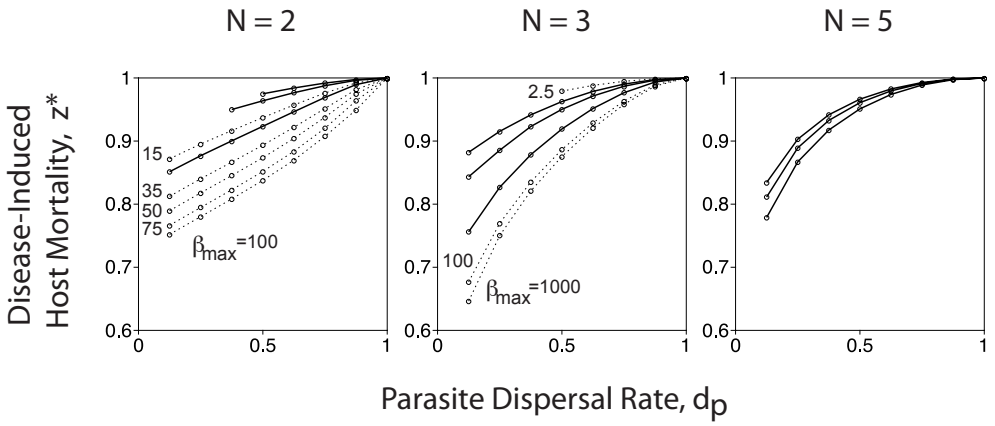

Fig. S1. Relationship between stable level of disease-induced host mortality (i.e. ES virulence, z^*) and pathogen dispersal rate, d_p as both patch carrying capacity (N) and maximum disease transmissibility ($\beta_{max} = \lim_{x\to\infty}\beta(x)$) vary. Remaining parameters were $\mu = 1$, $d_h = 0.3$ and $b = 3$, and $\beta(x)$ was assumed to take the form, $\beta_{max}\, x/(1+x)$ (see text for parameter definitions). From top to bottom β_{max} = 5, 7.5, and 20, respectively. Results for additional values of β_{max} are presented as dashed lines (values are as indicated). The qualitative pattern illustrated in panels corresponding to N = 3, 5 was also identified for N = 10, however the effect was too small to be seen easily along side other plots.

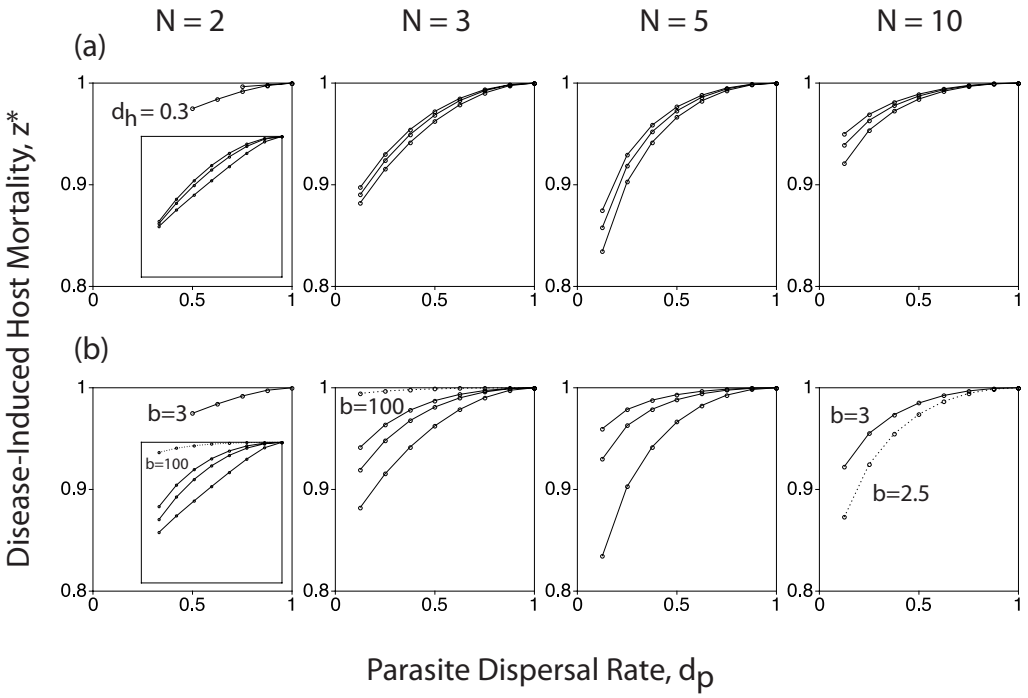

Fig. S2. Relationship between stable level of disease-induced host mortality (i.e. ES virulence, z^*) and pathogen dispersal rate, d_p as both patch carrying capacity (N) and host life-history parameters vary. Remaining parameters were μ= 1 and $\beta(x) = 5x/(1+x)$ (see text for definitions). (a) From top to bottom host dispersal takes values d_h = 0.9, 0.6, and 0.3 respectively under the assumption that b = 3. (b) From top to bottom host dispersal takes values b = 9, 6, and 3 respectively under the assumption that d_h = 0.3. Results for additional values of b are presented as dashed lines in (b) (values are as indicated). To better elucidate the effect of changes in host life-history parameters for the case N = 2, numerical analyses were also carried out assuming $\beta(x) = 15x/(1+x)$. These additional results are presented as inset figures whose axes display the same range of d_p and z^* values used in the main panels.