TITLE: The Importance of Who Infects Whom: The Evolution of Diversity in Host Resistance to Infectious Disease

Supplementary Information

In the main paper we focus on a model framework that considers n host types and represents the dynamics of susceptible hosts of type h, X_h, and infected hosts of type h, Y_h, with the following equations.

$$dX_h/dt = a_h X_h - q_h H X_h - b_h X_h - \sum_k \beta_{hk} X_h Y_k + \gamma_h Y_h \qquad (S1)$$

$$dY_h/dt = \sum_k \beta_{hk} X_h Y_k - \Gamma_h Y_h \qquad (S2)$$

where $h, k = 1,...,n$, $H = \sum_h X_h + \sum_h Y_h$ and $\Gamma_h = \alpha_h + b_h + \gamma_h$.

Under the assumption that the evolving life history parameters are the host birth rate and disease transmission (as in the main text) then the fitness, $\bar{s}$, of a mutant (subscript m) attempting to invade a resident (subscript r) population at equilibrium is given by determining the largest eigenvalue of the mutant submatrix, J_{mut}, of the Jacobian of (S1) and (S2). This Jacobian is given by

$$J_{mut} = \begin{pmatrix} a_m - b - qH_r - \beta_{mr}Y_r & \gamma \\ \beta_{mr}Y_r & -(\alpha + b + \gamma) \end{pmatrix} \qquad (S3)$$

We use the Next Generation Theorem to compute the mutant host's invasion fitness by finding the largest eigenvalue of J_{mut} (see Hurford et al. 2010). In particular we can set $J_{mut} = F - V$ where F is a matrix of all terms involving creation of new hosts (i.e. births) and V is a matrix of all terms involving removal of hosts (i.e. death) and transition of hosts between states, such that,

$$F = \begin{pmatrix} a_m - qH_r & 0 \\ 0 & 0 \end{pmatrix}, \quad V = \begin{pmatrix} b + \beta_{mr}Y_r & -\gamma \\ -\beta_{mr}Y & (\alpha + b + \gamma) \end{pmatrix}$$

Given this decomposition, then it can be shown that $\bar{s} = \rho(FV^{-1}) - 1$ is equivalent to the fitness of the mutant host given by the dominant eigenvalue of J_{mut} (see Hurford et al. 2010). We find that

$$\bar{s}(a_m, \beta_{mr}, a_r, \beta_{rr}) = \frac{(a_m - qH)(\alpha + b + \gamma)}{b(\alpha + b + \gamma) + \beta_{mr}Y_r(\alpha + b)} - 1 \qquad (S4)$$

For the mutant to invade we require that $\bar{s} > 0$. To simplify analytical calculations we use a sign equivalent proxy of the fitness (found by rearranging $\bar{s} > 0$ in (S4) and which gives equivalent results for the singular strategy, ES and CS conditions when considering single species evolution). This is

$$s(a_m, \beta_{mr}, a_r, \beta_{rr}) = a_m - qH - b + \frac{\beta_{mr}Y_r(\alpha + b)}{(\alpha + b + \gamma)} \qquad (S5)$$

with the condition that the mutant can invade if $s > 0$ in (S5).

In the main text we illustrate the community dynamics and adaptive dynamics analysis that examines the simplifying case where $\beta_{hk} = \beta_h$ (which we will refer to as **case 1**). In this document we present the full analysis for all the cases discussed in the main paper.

Case 1: Let us consider the case where $\beta_{hk} = \beta_h$ such that susceptibility to infection is host specific but transmissibility is the same for all (infected) hosts.

We can extract the β_h from the summation terms in equations (S1) and (S2) and defining $Y = \sum_k Y_k$ we find from (S2) that at equilibrium,

$$(\beta_h / \Gamma_h) X_h Y - Y_h = 0 \quad h = 1,...n \tag{S6}$$

and further using equation (S6) we can substitute for Y_h in (S1) at equilibrium to give

$$X_h \left\{ a_h - q_h H - b_h - ((\Gamma_h - \gamma_h)/\Gamma_h) \beta_h Y \right\} = 0 \quad h = 1,...n. \tag{S7}$$

Equation (S7) can be satisfied if $X_h = 0$ but we are interested in solutions when $X_h \neq 0$. By summing equation (S6) over h and using the constraints associated with (S1) and (S2) we can simplify the equations that determine the equilibrium densities (when $\beta_{hk} = \beta_h$ and when $X_1...X_l \neq 0$) to

$$\left(a_h - q_h H - b_h - \left(\beta_h (\Gamma_h - \gamma_h)/\Gamma_h \right) Y \right) = 0 \quad h = 1,...,l. \tag{S8}$$

$$\sum_h (\beta_h / \Gamma_h) X_h - 1 = 0 \tag{S9}$$

$$H - \sum_h X_h + Y = 0. \tag{S10}$$

These are equation (3)-(5) in the main text. We wish to solve equation (S8)-(S10) to find solutions $X_1,...,X_l,Y,H$. We therefore systematically examine equations (S8)-(S10) for different values of l to determine where there are consistent solutions. Thus generically from equation (S8): (i) for $l > 2$ there are more than two equations in two unknowns, Y,H, and so there are no solutions (ii) for $l = 2$ there is a unique solution for Y,H, (iii) for $l = 1$ there are many solutions for Y,H; from equation (S9): (i) for $l = 1$ there is a unique solution for X_1, (ii) for $l > 1$ there is one equation and more than one unknown, $X_1,...,X_l$ and so there are many solutions; and from equation (S10): (i) for $l \geq 1$ there is one equation in many unknowns $X_1,...,X_l,Y,H$ and so there are many solutions. Generically, (S8), (S9) and (S10) has a unique solution X_1,Y,H when $l = 1$ (S8iii, S9i and S10i) or a unique solution X_1,X_2,Y,H when $l = 2$ (S8ii, S9ii and S10i) and there are no solutions for $l > 2$ as equation (S8) cannot be satisfied. These conditions are necessary and sufficient for equilibrium solutions (but not sufficient for feasibility and stability). The key insight from the community dynamics is that no more than two host types can coexist and that therefore the maximum level of diversity that can occur is dimorphism.

The results highlighted by the community dynamics can be examined in the full evolutionary model by undertaking an adaptive dynamic (AD) analysis. For the general case (equations S1 and S2) it can be shown that a proxy for the fitness of a rare mutant (with parameters denoted with subscript m) attempting to invade a resident (subscript r) is as follows:

$$s(a_m, \beta_{mr}, a_r, \beta_{rr}) = a_m - qH_r - b - \beta_{mr}Y_r\left(\frac{\alpha+b}{\alpha+b+\gamma}\right) \qquad \text{(S11)}$$

Under the assumptions that $\beta_{hk} = \beta_h$ and that there is a trade-off $a_i = f(\beta_i)$ then an evolutionary singular strategy, β^*, occurs when the fitness gradient is zero

$$\left.\frac{\partial s}{\partial \beta_m}\right|_{\beta_m = \beta_r} = 0 \quad \Rightarrow \quad f'(\beta^*) = Y^*\left(\frac{\alpha+b}{\alpha+b+\gamma}\right) \qquad \text{(S12)}$$

Here, Y^* is the equilibrium value of Y when $\beta_h = \beta^*$. The evolutionary behaviour at the singular point is determined by the following conditions:

Evolutionary Stability (ES): $\left.\dfrac{\partial^2 s}{\partial \beta_m^2}\right|_{\beta^*} < 0$

$$\Rightarrow \quad f''(\beta^*) < 0 \qquad \text{(S13)}$$

Convergent Stability (CS): $\left.\dfrac{\partial^2 s}{\partial \beta_r^2}\right|_{\beta^*} - \left.\dfrac{\partial^2 s}{\partial \beta_m^2}\right|_{\beta^*} > 0$

$$\Rightarrow \quad f''(\beta^*) < \frac{q(\alpha+b)(\alpha+b+\gamma)}{(\beta^*)^2\left(q(\alpha+b+\gamma) + \beta^*(\alpha+b)\right)} = \Phi \qquad \text{(S14)}$$

Evolutionary branching occurs when a singular point is convergence stable (attracting) but evolutionarily unstable (invadable), and provided the resident and mutant types are mutually invadable at the singular point (which is guaranteed in evolutionary systems if the conditions for branching hold). In this model evolutionary branching leading to dimorphism can occur for trade-offs that satisfy $0 < f''(\beta^*) < \Phi$ (as here the system is not ES but is CS in equations (S10) & (S11) respectively).

To examine whether further branching (further diversity) can occur we consider a dimorphic resident population and examine whether branching can occur from either resident host type. After branching, the dimorphic population will follow a unique evolutionary trajectory until either it reaches a co-singular point (i.e. a singular point for both types) or the maximum/minimum limits of evolution. Consider a rare mutant with parameter β_m attempting to invade a resident population with parameters β_{r1} and β_{r2}. Assuming the mutant type is close to β_{r1} the proxy for fitness can be written as

$$s(a_m, \beta_m, a_{r1}, \beta_{r1}) = f(\beta_m) - qH_r - b - \beta_m\left(Y_{r1} + Y_{r2}\right)\left(\frac{\alpha+b}{\alpha+b+\gamma}\right) \qquad \text{(S15)}$$

Assuming both types are at the co-singular point, then the singular point for type β_{r1} occurs at

$$\left.\frac{\partial s}{\partial \beta_m}\right|_{\beta_m = \beta_{r1}} = 0 \quad \Rightarrow \quad f'(\beta_m) = \left(Y_{r1} + Y_{r2}\right)\left(\frac{\alpha+b}{\alpha+b+\gamma}\right)\Bigg|_{\beta_m = \beta_{r1}} \qquad \text{(S16)}$$

For evolutionary stability we again have

ES: $\Rightarrow \quad f''(\beta_m)\big|_{\beta_m = \beta_{r1} = \beta^*} < 0 \qquad \text{(S17)}$

The condition for convergence stability is now more complex as it depends on the evolutionary trajectory of both host types. However, we know that a requirement for

evolutionary branching (even at a convergent but evolutionarily unstable singular point) is mutual invadability. It can be shown that,

$$\text{Mutual Invadability (MI): } \left.\frac{\partial^2 s}{\partial \beta_m \partial \beta_{r1}}\right|_{\beta_m = \beta_{r1}} = 0 \qquad (S18)$$

As such host types are unable to build up either side of the singular point, and further evolutionary branching will never occur. By symmetry this will be exactly the same for a mutant close to β_{r2}. Therefore further branching from either of the host types cannot occur and both the adaptive dynamics and community dynamics reach the same conclusion. These results are supported by simulations of the evolutionary process.

Case 2: Let us now consider the case where $\beta_{hk} = \beta_h \mu_k$ such that transmission is composed of the hosts susceptibility to infection and the transmissibility of other hosts in a multiplicative manner.

The community dynamics analysis for the case $\beta_{hk} = \beta_h \mu_k$ proceeds as follows. The β_h term can once again be extracted from the summation terms in (S1) and (S2) but the μ_k term cannot and so we define $Z = \sum_k \mu_k Y_k$. Equations (3)-(5) in the main text are replaced by the following equations and we seek solutions where $X_1 ... X_l \neq 0$

$$\left(r_h - q_h H - \left(\beta_h \left(\Gamma_h - \gamma_h\right)/\Gamma_h\right)Z\right) = 0 \quad h = 1,...,l \qquad (S19)$$

$$Z - \sum_k \mu_k Y_k = 0 \qquad (S20)$$

$$\sum_h \left(\beta_h \mu_h/\Gamma_h\right)X_h - 1 = 0 \qquad (S21)$$

$$H - \sum_h X_h + Z\sum_h \left(\beta_h/\Gamma_h\right)X_h = 0 \qquad (S22)$$

The polynomial form of equation (S22) means that to make progress we need to assume that a solution exists which allows us to make a linear approximation of (S22) near to this solution. Then, in order, the separate parts provide l linear equations in 2 unknowns, Z, H 1 linear equation in $l+1$ unknowns $Y_1,...,Y_l, Z$, 1 linear equation in l unknowns $X_1,...,X_l$ and 1 linear equation in $l+3$ unknowns $X_1,...,X_l, Y, Z, H$. Thus generically:

From (S16):
  i.   For $l > 2$ there are no solutions for Z, H
  ii. For $l = 2$ there is a unique solution for Z, H
  iii. For $l = 1$ there are many solutions for Z, H

From (S17):
  i.   For $l \geq 1$ there are many solutions for $Y_1,...,Y_l, Z$

From (S18):
  i.   For $l = 1$ there is a unique solution for $X_1,...,X_l$
  ii. For $l > 1$ there are many solutions for $X_1,...,X_l$

From (S19)
  i.   For $l \geq 1$ there are many solutions for $X_1,...,X_l, Z, H$

Consistent solutions are only possible when $l = 1$ (S19(iii) and S21(i)) or $l = 2$ (S19(ii) and S21(ii)). Generically, these conditions are necessary but not sufficient for equilibrium solutions. Under this set-up no more than two host strains can coexist.

For the adaptive dynamics analysis we assume a trade-off $a_i = f(\beta_i)$ (as in case 1) but we also assume that host transmissibility is linked to the level of host resistance such that $\mu_j = g(\beta_j)$. The fitness expression is therefore

$$s(a_m, \beta_m, a_r, \beta_r) = f(\beta_m) - b - qH_r - \beta_m g(\beta_r) Y_r \left(\frac{\alpha + b}{\alpha + b + \gamma} \right) \qquad \text{(S23)}$$

A singular strategy, β^*, occurs when

$$\left. \frac{\partial s}{\partial \beta_m} \right|_{\beta_m = \beta_r} = 0 \quad \Rightarrow \quad f'(\beta^*) = g(\beta^*) Y^* \left(\frac{\alpha + b}{\alpha + b + \gamma} \right) \qquad \text{(S24)}$$

The singular strategy is ES if $f''(\beta^*) < 0$ \qquad (S25)

The singular strategy is CS if $f''(\beta^*) < \Phi_1 \left(1 + g'(\beta^*) \Phi_2 \right)$ \qquad (S26)

where $\Phi_1 = \dfrac{q(\alpha + b)(\alpha + b + \gamma)}{(\beta^*)^2 \left(q(\alpha + b + \gamma) + \beta^* g(\beta^*)(\alpha + b) \right)}$ and $\Phi_2 = \dfrac{(\alpha + f(\beta^*)(\beta^*)^2}{q(\alpha + b + \gamma) + \beta^* g(\beta^*)(\alpha + b)}$.

Note, that $\Phi_1, \Phi_2 > 0$. If, as in case 1, $g(\beta) = 1$ then $g'(\beta) = 0$ and the CS condition becomes $f''(\beta^*) < \Phi_1$ (note also that here $\Phi_1 = \Phi$ and we recover the CS condition for case 1 shown in equation (S14)). If $g'(\beta) > 0$ then transmissibility and susceptibility are correlated in a positive manner such that types that are most likely to be infected (high β) are also most likely to infect (high μ). In this scenario the region of branching is increased. If $g'(\beta) < 0$ then transmissibility and susceptibility are correlated in a negative manner such that types that are least likely to be infected (low β) are also most likely to infect (high μ). In this scenario the region of branching is decreased and depending on the magnitude of $g'(\beta)$ may disappear and be replaced by a region where the singular strategy exhibits garden of eden behaviour (specifically if $g'(\beta) < -1/\Phi_2$).

As in case 1 we can examine whether further branching can occur. We can examine the success of a rare mutant attempting to invade a dimorphic resident population. As in the main text it can be shown analytically that the condition for mutual invadability (MI) is zero at each of the resident strains composing the dimorphic resident population. Therefore host strains are unable to build up either side of a dimorphic singular point, and further evolutionary branching cannot occur. These results are supported by simulations of the evolutionary process.

Case 3: Let us now consider the community dynamics when β_{hk} is not simplified. It can be shown that the equilibrium points (for $X_1...X_l \neq 0$) are solutions of

$$\left(r_h - q_h H - \left((\Gamma_h - \gamma_h)/\Gamma_h\right)\sum_k \beta_{hk} Y_k\right) = 0 \quad h = 1,...,l \tag{S27}$$

$$\sum_k (\beta_{hk}/\Gamma_h) X_h Y_k - Y_h = 0 \quad h = 1,...l \tag{S28}$$

$$H - \sum_h X_h + \sum_h Y_h = 0. \tag{S29}$$

The polynomial form of equation (S28) means that to make progress we need to assume that a solution exists which allows us to make a linear approximation of (S28) near to this solution. In order, the separate parts provide l linear equations in $l+1$ unknowns Y, H; l linear equation in $2l$ unknowns $X_1,..., X_l, Y_1,..., Y_l$ and 1 linear equation in $2l+1$ unknown $X_1,..., X_l, Y_1,..., Y_l, H$. Thus generically each part has multiple solutions with codimensions respectively l, l and 1 and the linear system has a unique solution. Generically, there are no conditions on l necessary for equilibrium solutions (but feasibility and stability is not determined). This means there is no restriction on the coexistence of different host strains and so polymorphism is possible.

In the main text we consider an example of a transmission term where transmission depends on both host susceptibility, β_h, and the interaction between the susceptible, h, and infectious, k, host types such that transmissibility is greatest between similar host types such that

$$\beta_{hk} = \beta_h \left[1 - c + c \exp\left(-\left(\frac{h-k}{w}\right)^2\right)\right].$$ Here transmission depends on the 'relatedness' of

susceptible and infectious host types and is greater between similar types. As in case 1 we assume a trade-off $a_h = f(\beta_h)$. The fitness expression is

$$s(a_m, \beta_{mr}, a_r, \beta_{rr}) = a_m - b - qH_r - \beta_{mr} Y_r \left(\frac{\alpha + b}{\alpha + b + \gamma}\right) \tag{S30}$$

A singular strategy, β^*, occurs when

$$\left.\frac{\partial s}{\partial \beta_m}\right|_{\beta_m = \beta_r} = 0 \quad \Rightarrow \quad f'(\beta^*) = Y^* \left(\frac{\alpha + b}{\alpha + b + \gamma}\right) \tag{S31}$$

The singular strategy is ES if $f''(\beta^*) < -\frac{2\beta^* c}{w^2} Y^* \left(\frac{\alpha + b}{\alpha + b + \gamma}\right)$ (S32)

The singular strategy is CS if $f''(\beta^*) < \frac{q(\alpha + b)(\alpha + b + \gamma)}{(\beta^*)^2 \left(q(\alpha + b + \gamma) + \beta^*(\alpha + b)\right)} = \Phi$ (S33)

Note the CS condition is that same as that for case 1 (equation S14) and the right hand side of the ES condition becomes zero if $c = 0$ (and here case 3 is equivalent to case 1). The region

that supports evolutionary branching increases as c increases or as w decreases. These results are supported by simulations of the evolutionary process.

Simulations: In the simulations shown in the figures, the population dynamics were numerically solved for a fixed time (t_a) according to equations (1) and (2) initially with a monomorphic population. Therefore we initially have a single host, $h = r$, that has susceptibility to infection β_r (since h and β_h are linked) and a birth rate a_r (since a_h is linked to β_h through the trade-off). A mutant type is generated as a small deviation around the current resident trait (and has either a slightly higher or slightly lower trait value) and therefore a mutant type, $h = m$ (where m is close to r, and therefore β_m is close to β_r) is introduced at low density. The population dynamics were then solved for a further time t_a with types whose population density fell below a (low) threshold considered extinct and removed before considering new mutations. This way the trait h (and β_h and a_h) can evolve and evolutionary branching can lead to the coexistence of dimorphisms and polymorphisms. When more than one type coexists the choice of current type from which to mutate depends on its relative population density.

References
Hurford, A., Cownden, D. and Day, T. (2010) Next-generation tools for evolutionary invasion analyses. *J. R. Soc. Interface*, 7, 561-571.