Supporting Information

Rudge et al. 10.1073/pnas.1221509110

SI Text

Bayesian Estimation of the True Prevalence of Infection. As is common with diagnostic tools used in parasitological surveys, none of the tests used in the field surveys can be considered gold standards. Serological tests are generally known to overestimate the population prevalence of infection, because they may detect antibodies from a previous infection that has now cleared. Stool examinations, however, generally underestimate the prevalence of infection because of a relatively low sensitivity, particularly for detecting light infections (1–4). A Bayesian model was, therefore, applied to estimate the true prevalence of *Schistosoma japonicum* infection among each definitive host species and snail intermediate hosts in each village using methods from previous studies (2, 3, 5). Table S1 presents the census and parasitological data for mammalian definitive hosts, and Table S2 presents the census and parasitological data for the snail intermediate host. The model is based on the premise that the probability (p) of any single test on any given animal being positive can be expressed as $p = \pi\, S + (1-\pi)\,(1-C)$, where π is the true prevalence of infection in the animal population under study and S and C are the sensitivity and specificity of the test, respectively. (Sensitivity measures the proportions of true positives that are correctly identified as such by the test, and specificity measures the proportions of true negatives that are correctly identified.) Prior distributions were first constructed for the sensitivity and specificity of each test (Table S3) and the prevalence of infection for each host population (Table S4). By combining these prior distributions with the data through the likelihood function, a posterior distribution for the true prevalence of infection was derived.

Ranges for the sensitivity and specificity of each test were derived from a review of the relevant literature and expert opinion (Table S3). Very little data could be found on the test properties of the miracidia hatching test, particularly for nonhuman animals, and therefore, the sensitivity range for this test was set wide at 50–95% for all definitive host species. This range is comparable with the 95% credible intervals previously estimated for the sensitivity of an alternative stool examination method, the sedimentation technique, when used on a single stool sample to detect *S. japonicum* infection in various nonhuman hosts in the Philippines (5). The sensitivity of the hatching test used in the present study was considered unlikely to be less than 50%; the test is, if anything, likely to be more sensitive than the sedimentation technique, because a higher amount of feces is examined in the former. Because *S. japonicum* miracidia are easily identifiable under a microscope, the specificity of the hatching test was assumed to be known at 100%. Similarly, the specificity of the crushing method for identifying snail infections was assumed to be 100%.

The ranges derived for unknown sensitivities and specificities for each diagnostic test were used to inform the prior distributions for these parameters in the Bayesian model. These priors were assumed to follow a β-distribution, with the mean (μ) of the β-distribution matched to the center of the range and the SD (σ) matched to one-quarter of the total range. These conditions define the two coefficients of the β-distribution, α and β, as follows (2, 3):

$$\alpha = \mu\left(\frac{(1-\mu)\mu}{\sigma^2} - 1\right) \qquad \text{[S1]}$$

and

$$\beta = (1-\mu)\left(\frac{(1-\mu)\mu}{\sigma^2} - 1\right). \qquad \text{[S2]}$$

The β-family of distributions was chosen, because its region of positive density is between zero and one (and therefore, it matches the possible range for sensitivity, specificity, and prevalence). Also, it is very flexible, in that a wide variety of possible shapes can be defined by using different values of α and β (3); β-prior distributions were, therefore, also assigned to the prevalence of infection for each host species. For humans and livestock, these distributions were informed by previous national surveillance data collected in Anhui in 2005 and data from Wang et al. (6) (also collected in Anhui), with the mean and SD of the β-prior matched with the mean and SD of the observed prevalence across villages (Table S4). Because very few or no dogs, cats, or rodents were sampled in the 2005 national surveillance survey, the prevalence for each of these species was given an uninformative prior of β (1, 1) (equivalent to a uniform distribution from zero to one).

The model was run in WinBUGS (WinBUGS version 1.4; www.mrc-bsu.cam.ac.uk/bugs/) using Markov Chain Monte Carlo simulations with a burn-in phase of 10,000 iterations followed by 20,000 iterations for inference. Separate analyses were carried out for each host species in each village, such that the priors for prevalence, sensitivity, and specificity were modeled independently between species and villages. Because of a low sample size for some species in some villages and to estimate the prevalence in each species in an average village of each habitat type, the model was also run for data that had been pooled across all three villages in each region. The posterior distributions for the true prevalence of infection in each species are given in Table S5.

Choice of Transmission Model Framework. A prevalence model framework, based on the model by Barbour (7), was adopted instead of an intensity model framework (which tracks the number of adult parasites among definitive hosts) for a number of reasons. First, although intensity models, such as those models based on a framework initially devised by Macdonald (8), have, in principle, the advantage of capturing heterogeneities in worm burden among hosts, they are difficult to parameterize. Intensity models of multihost systems can be particularly challenging to parameterize, requiring accurate estimation of worm burden in all possible hosts involved in transmission. In low-transmission settings, measurements of infection intensity (which are, by necessity, indirect and reliant on egg counts) suffer from issues of poor diagnostic sensitivity, which was particularly true of egg counts in the human population sampled in our villages. Measuring infection intensity in other possible reservoirs is also fraught with difficulties. Lethal sampling of nonhuman hosts for quantification of worm burden is often not possible, and functional relationships between adult worm burden and rates of egg production, for each putative reservoir, are largely unknown.

Furthermore, according to the relationship between infection prevalence and load that derives from assuming a negative binomial distribution of parasites per host, this relationship becomes more linearly proportional in low-transmission settings, with changes in intensity being better reflected by changes in prevalence in high-transmission settings (9).

Finally, as shown by Barbour (7), there are other shortcomings inherent to the Macdonald (8) type of models (based on infection intensity), resulting in underestimates of the basic reproduction number (R_0), even after adjusting for heterogeneous exposure rates. Subsequently, the Barbour (7) model has been successfully

adapted to the study of the transmission dynamics of *S. japonicum*, incorporating nonhuman mammalian reservoirs in both China (10, 11) and the Philippines (12, 13), and also a multihost transmission model of *S. mekongi*, a sister taxon of *S. japonicum*, in Cambodia (14).

Estimation of Transmission Rates. Transmission rates of the model were estimated from prevalence data, assuming that these represent steady state values. By setting the left-hand sides of Eqs. **1** and **2** to zero, formulas for transmission rates a_i and b_i are derived,

$$a_i = \frac{g_i P_i^*}{\Delta\left(1 - P_i^*\right)\sum_j \omega_{ij} y_j^*} \quad \text{[S3]}$$

and

$$b_i = \frac{\gamma \Delta y_i^*}{\phi_i P_i^* \left(1 - Y^*\right)}, \quad \text{[S4]}$$

where P_i^* and y_i^* denote the steady state values for prevalence levels in definitive host species i and snail intermediate hosts infected by host species i, respectively, and Y^* denotes the overall steady state prevalence in snails. Values (or at least plausible ranges of values) for parameters P_i^*, Y^*, Δ, γ, ϕ_i, and g_i can be estimated from local parasitological survey data and data from the literature, and values for ω_{ij} were allowed to vary between 0.01 and 1 as described below (*SI Text*, section 5). However, values of y_i^* cannot be directly inferred from local parasitological data, because snail surveys can only estimate Y^* and cannot differentiate snail infections according the species of definitive host that gave rise to the infection, leaving Eqs. **S3** and **S4** underspecified. To overcome this problem, b_i for each host species was assumed to be proportional to the average rate of egg excretion of an infected individual of that species, such as in the work by Williams et al. (10), and therefore,

$$b_i = \varepsilon_i \theta_i \eta, \quad \text{[S5]}$$

where ε_i is the average number of eggs per gram of feces excreted by an infected host of species i and θ_i is the amount of feces excreted in grams per unit time by a host of species i, both of which are directly measurable. The parameter-η is, therefore, a constant across all species within a given village and can be interpreted as the probability that an egg excreted by any definitive host successfully causes a snail infection. Using Eq. **S5**, the rate of change in the overall prevalence among snails can be expressed as

$$\frac{dY}{dt} = \eta \sum_i (\varepsilon_i \theta_i \phi_i P_i) \frac{1}{\Delta}(1 - Y) - \gamma Y. \quad \text{[S6]}$$

Eq. **S6** shows that the relative contribution of each definitive host species i to the overall prevalence among snails, Y, is proportional to its rate of egg excretion, $\varepsilon_i \theta_i$, population density, φ_i, and prevalence of infection, P_i. Parameter-η can then be estimated by setting $\frac{dY}{dt}$ to zero:

$$\eta = \frac{\Delta \gamma Y^*}{\sum_i \left(\varepsilon_i \theta_i \phi_i P_i^*\right)\left(1 - Y^*\right)}. \quad \text{[S7]}$$

Eq. **S5** is then fully specified to estimate b_i, which in turn, allows y_i^* to be estimated by rearranging Eq. **S4**. Eq. **S3** is then fully specified to estimate a_i.

Estimation of the Basic Reproduction Number R_0. Adopting the notation from the work by Roberts and Heesterbeek (15), if there are n definitive host species, we can construct an $n \times n$ next generation matrix, K. In the present study, the term generation refers to a transmission event from a definitive host to a snail intermediate host to another definitive host. Each element, k_{ij}, of matrix K is, therefore, the expected number of definitive hosts of species i that would be infected (by the snail intermediate host population) by a primary infected definitive host of species j in a susceptible population. Thus, k_{ij} is similar in concept to R_0. From the system described in Eqs. **1** and **2**,

$$k_{ij} = \frac{\omega_{ij} a_i b_j \phi_j}{g_j \gamma}. \quad \text{[S8]}$$

Eq. **S8** is somewhat intuitive, because new infections in snails caused by an infected definitive host of species j arise at rate $b_j \phi_j$ during $1/g_j$, whereas new infections in host species i caused by these infected snails arise at rate $\omega_{ij} a_i$ during $1/\gamma$.

The elements in the diagonal of matrix K (i.e., when $i = j$) can be interpreted as the basic reproduction number of the infection within species i denoted by $R_0^{(i)}$. Because ω_{ij} is set to one when $i = j$ (in the main text), the formula for $R_0^{(i)}$ is

$$R_0^{(i)} = \frac{a_i b_i \phi_i}{g_i \gamma}. \quad \text{[S9]}$$

Therefore, in a simple example of two definitive host species, humans (H) and cattle (C), the next generation matrix K is defined as

$$K = \begin{pmatrix} k_{HH} & k_{HC} \\ k_{CH} & k_{CC} \end{pmatrix} = \begin{pmatrix} \frac{a_H b_H \phi_H}{g_H \gamma} & \frac{\omega_{HC} a_H b_C \phi_C}{g_C \gamma} \\ \frac{\omega_{CH} a_C b_H \phi_H}{g_H \gamma} & \frac{a_C b_C \phi_C}{g_C \gamma} \end{pmatrix}. \quad \text{[S10]}$$

The overall basic reproduction number for the system, R_0, is calculated as the dominant eigenvalue (spectral radius) of the next generation matrix, K, written as $R_0 = \rho(K)$ (15, 16).

It is worth noting that, when all values for $\omega_{ij} = 1$ [i.e., with homogenous mixing between definitive host species as in the original model by Barbour (7)], $\rho(K)$ is equal to the sum of $R_0^{(i)}$ across all i, and therefore,

$$R_0 = \sum_i \frac{a_i b_i \phi_i}{g_i \gamma} = \sum_i R_0^{(i)}. \quad \text{[S11]}$$

However, with heterogeneous mixing across host species, Eq. **S11** does not hold, and it is more difficult to obtain an analytical solution for R_0. The dominant eigenvalue of K was, therefore, computed numerically (using the eigen function in R version 2.15) to calculate the overall R_0.

Transmission Model Parameter Values and Distributions. Parameter values and their sources are presented in Table S6. The rates of fecal excretion, θ_i, for humans and domesticated animals were assigned triangular distributions based on the mean, minimum, and maximum values reported for each species in the work by Wang et al. (6). The rate of fecal excretion for rodents was assigned a uniform distribution between 1 and 2 g d^{-1} (17). The average numbers of eggs excreted per gram of feces, ε_i, for an infected individual of each host species were given uniform distributions between ranges informed by data collected in the field (6, 18).

Infected snail mortality, γ, was given a uniform distribution based on the reciprocal of the range of average life expectancy (30–158 d) reported for infected *Oncomelania hupensis* snails (19–21). The recovery rate for human infections, g_H, was assigned a triangular distribution based on an estimated average duration of

infection of 1 y (range = 0.5–1.5 y) because of the annual rate of drug administration and the fact that praziquantel is adulticidal against schistosomes. For infections in domesticated animals, the recovery rates were given uniform distributions based on an estimated range of 2–5 y for the duration of infection in these animals. This range covers the 3- to 5-y average lifespan reported for adult worms (20) but also allows for the possibility that the duration of infection is shorter than this rate because f a potentially short life expectancy of these animals (13). Because rodents in the wild are short-lived relative to the adult parasites, they were assumed to die before recovery at a rate reciprocal to an assumed average lifespan of 1–2 y (22, 23).

The population densities of domesticated animals in each village and region were given single-point estimates based on the local survey data (Table S1). The density of humans, ϕ_H, in each village was given a uniform distribution, with a minimum value based on the number of individuals examined using the indirect hemagglutination assay (IHA) in the parasitological survey and a maximum value based on the total registered human population. This range was deemed to be sensible, because as reported by local village leaders, many of the registered population may not actually reside within the village (for instance, having relocated to nearby towns for employment); therefore, they would not play a role in local transmission. Because the density of rodents, ϕ_R, within each village could not be measured, it was given a wide uniform distribution. In hilly villages, where the trapping success of rodents was relatively high, ϕ_R was given a range between 10 and 100 rodents per hectare (ha^{-1}) snail habitat. This range covers the majority of rodent densities reported by various studies across a range of ecological settings (24–28). In the marshland region, the trapping success of rodents per trap-day was only 23% of the rate in the hilly region (Table S1). The range for the rodent density in this region was, therefore, scaled down proportionately to the relative trapping success rate, and therefore, ϕ_R in marshland villages was given a uniform distribution between 2.3 and 23 individuals ha^{-1} snail habitat.

It was not possible to infer from any data the interhost species values for ω_{ij}, which effectively determine the level of spatial overlap or mixing between hosts of species i and j relative to the level within a given host species (i.e., when $i = j$, for which ω_{ij} was set to unity). Therefore, interhost species values of ω_{ij} were allowed to vary between a wide range of 0.01 and 1. A value of 0.01 for ω_{ij} could be interpreted as meaning that host species i shares only 1% of its snail-inhabited water contact sites with host species j. Choosing this value as the minimum value for ω_{ij} was somewhat arbitrary, although for anything less than this value, one might expect to see stronger *S. japonicum* population structuring among host species than was observed in our previous molecular data (29). The scenario of $\omega_{ij} = 0$ was not considered, because the generally very high levels of parasite gene flow observed between host species rule out the possibility of no *S. japonicum* transmission between host species.

For steady state prevalence estimates among definitive hosts (P_i^*) and snails (Y^*) in each village and region, rather than sampling from parametric statistical distributions, we sampled from the posterior distributions derived from Bayesian analysis of local parasitological survey data (Table S5). These distributions account for both statistical uncertainty (because of limited numbers of individuals sampled) and experimental uncertainty (because of imperfect diagnostic test properties) in the steady state prevalence estimates.

To sample the parameter space efficiently, the Latin Hypercube Sampling (LHS) method was used (30, 31). For this method, the probability distribution of each of κ-uncertain parameters is divided into n equally probable intervals. A value for each parameter is then chosen at random for each interval, resulting in n sets of parameters, which then allows n simulations to be carried out, each using a different set of parameters. In this study, a total of 53 input parameters was allowed to vary in the uncertainty analysis. For each village and region, 500 sets of parameters were generated using LHS, and thus, 500 simulations were performed. With 53 uncertain parameters, this number of simulations easily satisfies the criterion that, for κ-uncertain parameters, at least $4\kappa/3$ simulations should be performed when using LHS to sample the parameter space adequately (30–32). From each set of input parameters, transmission rates (a_i and b_i), $R_0^{(i)}$ values, and overall R_0 were calculated. Thus, for each village and region, 500 values were estimated for each of these outputs. The medians and 2.5 and 97.5 percentiles of these distributions of $R_0^{(i)}$ and R_0 were then calculated for each village and region. In addition, using the resulting a_i and b_i values, a longitudinal simulation was performed for each set of parameters using the system described by Eqs. **1** and **2** to track Pi and yi over time and thus, predict the impact of various intervention scenarios.

All model equations were implemented and solved in R version 2.15 (www.r-project.org/). The simecol package (33) was used to simulate transmission over time, using the Runge–Kutta fourth order algorithm for numerical integration.

Configuration of Multihost–Pathogen Communities. Fenton and Pedersen (34) recently proposed a conceptual framework to describe the configurations of multihost–pathogen communities. Based on a pathogen's within- and between-species transmission rates, this framework allows disease threats to be classified into one of four disease outcomes (with respect to a target host population) that can be summarized as follows:

i) Spillover, in which (*i*) the target host species cannot sustain transmission alone and (*ii*) transmission from the reservoir species to the target species is also low; therefore, although infections do occur in the target host species, they are transient (i.e., the pathogen does not seem to be endemic within the target species).

ii) Apparent multihost pathogen, in which the target host species cannot sustain transmission alone but transmission between species is high enough that infections are persistent (i.e., seem endemic) within the target host population. Note that this situation would also be described as spillover by Power and Mitchell (35).

iii) True multihost pathogen, in which both within- and between-species transmission rates are high; therefore, both species are considered maintenance hosts.

iv) Potentially emerging infectious disease, in which the pathogen can persist within the target host species, but the rate of transmission between species is so low that the target host is rarely exposed to the disease.

1. Yu JM, de Vlas SJ, Jiang QW, Gryseels B (2007) Comparison of the Kato-Katz technique, hatching test and indirect hemagglutination assay (IHA) for the diagnosis of *Schistosoma japonicum* infection in China. *Parasitol Int* 56(1):45–49.
2. Wang XH, Wu XH, Zhou XN (2006) Bayesian estimation of community prevalences of *Schistosoma japonicum* infection in China. *Int J Parasitol* 36(8):895–902.
3. Joseph L, Gyorkos TW, Coupal L (1995) Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. *Am J Epidemiol* 141(3):263–272.
4. Zhou YB, et al. (2007) Field comparison of immunodiagnostic and parasitological techniques for the detection of Schistosomiasis japonica in the People's Republic of China. *Am J Trop Med Hyg* 76(6):1138–1143.
5. Carabin H, et al. (2005) Estimating sensitivity and specificity of a faecal examination method for *Schistosoma japonicum* infection in cats, dogs, water buffaloes, pigs, and rats in Western Samar and Sorsogon Provinces, The Philippines. *Int J Parasitol* 35(14):1517–1524.
6. Wang TP, et al. (2005) Transmission of *Schistosoma japonicum* by humans and domestic animals in the Yangtze River valley, Anhui province, China. *Acta Trop* 96(2–3):198–204.
7. Barbour AD (1996) Modeling the transmission of schistosomiasis: An introductory view. *Am J Trop Med Hyg* 55(5 Suppl):135–143.
8. Macdonald G (1965) The dynamics of helminth infections, with special reference to schistosomes. *Trans R Soc Trop Med Hyg* 59(5):489–506.
9. Anderson RM (1987) Determinants of infection in human schistosomiasis. *Baillières Clin Trop Med Commun Dis* 2(2):279–300.

10. Williams GM, et al. (2002) Mathematical modelling of schistosomiasis japonica: Comparison of control strategies in the People's Republic of China. *Acta Trop* 82(2):253–262.
11. Gray DJ, Williams GM, Li Y, McManus DP (2008) Transmission dynamics of *Schistosoma japonicum* in the lakes and marshlands of China. *PLoS One* 3(12):e4058.
12. Riley S, et al. (2005) Estimating and modeling the dynamics of the intensity of infection with *Schistosoma japonicum* in villagers of leyte, Philippines. Part II: Intensity-specific transmission of *S. japonicum*. The schistosomiasis transmission and ecology project. *Am J Trop Med Hyg* 72(6):754–761.
13. Riley S, et al. (2008) Multi-host transmission dynamics of *Schistosoma japonicum* in Samar province, the Philippines. *PLoS Med* 5(1):e18.
14. Hisakane N, et al. (2008) The evaluation of control measures against *Schistosoma mekongi* in Cambodia by a mathematical model. *Parasitol Int* 57(3):379–385.
15. Roberts MG, Heesterbeek JA (2003) A new method for estimating the effort required to control an infectious disease. *Proc Biol Sci* 270(1522):1359–1364.
16. Diekmann O, Heesterbeek JA, Metz JA (1990) On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. *J Math Biol* 28(4):365–382.
17. Ebihara K, Shiraishi R, Okuma K (1998) Hydroxypropyl-modified potato starch increases fecal bile acid excretion in rats. *J Nutr* 128(5):848–854.
18. Lu DB, et al. (2010) Contrasting reservoirs for Schistosoma japonicum between marshland and hilly regions in Anhui, China—a two-year longitudinal parasitological survey. *Parasitology* 137(1):99–110.
19. Sun LP, et al. (2004) [Effect of *Schistosoma japonicum* infection on *Oncomelania hupensis*.]. *Zhongguo Xue Xi Chong Bing Fang Zhi Za Zhi* 16(4):265–268. Chinese.
20. Anderson RM, May RM (1991) *Infectious Diseases of Humans: Dynamics and Control* (Oxford Univ Press, Oxford).
21. Liang S, Spear RC, Seto E, Hubbard A, Qiu D (2005) A multi-group model of *Schistosoma japonicum* transmission dynamics and control: Model calibration and control prediction. *Trop Med Int Health* 10(3):263–278.
22. Jackson WB (1982) Norway rat and allies. *Wild Mammals of North America: Biology, Management, Economics*, eds Chapman JA, Feldhamer GA (Johns Hopkins University Press, Baltimore).
23. Nowak R, Paradiso J (1983) *Walker's Mammals of the World* (Johns Hopkins University Press, Baltimore, MD), 4th Ed.
24. Adler GH, Davis SL, Carvajal A (2003) Bots (Diptera: Oestridae) infesting a neotropical forest rodent, *Proechimys semispinosus* (Rodentia: Echimyidae), in Panama. *J Parasitol* 89(4):693–697.
25. Korn H (1987) Densities and biomasses of non-fossorial southern African Sanvanna rodents during the dry season. *Oecologia* 72(3):410–413.
26. Ku TY, Lin CC (1980) Abundance and distribution of field rodents in Taiwan. *Plant Prot Bull* 22(4):397–420.
27. Nakata K (1998) Regulation of reproduction in a natural population of the small Japanese field mouse, *Apodemus argenteus*. *Mammal Study* 23(1):19–30.
28. Sluydts V, Crespin L, Davis S, Lima M, Leirs H (2007) Survival and maturation rates of the African rodent, *Mastomys natalensis*: Density-dependence and rainfall. *Integr Zool* 2(4):220–232.
29. Rudge JW, et al. (2009) Parasite genetic differentiation by habitat type and host species: Molecular epidemiology of *Schistosoma japonicum* in hilly and marshland areas of Anhui Province, China. *Mol Ecol* 18(10):2134–2147.
30. Chan MS (1996) The consequences of uncertainty for the prediction of the effects of schistosomiasis control programmes. *Epidemiol Infect* 117(3):537–550.
31. Blower SM, Dowlatabadi H (1994) Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. *Int Stat Rev* 62(2): 229–243.
32. McKay MD, Conover WJ, Beckham RJ (1979) A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21(2):239–245.
33. Petzoldt T, Rinke K (2007) Simecol: An object-oriented framework for ecological modeling in R. *J Stat Softw* 22(9):1–31.
34. Fenton A, Pedersen AB (2005) Community epidemiology framework for classifying disease threats. *Emerg Infect Dis* 11(12):1815–1821.
35. Power AG, Mitchell CE (2004) Pathogen spillover in disease epidemics. *Am Nat* 164(Suppl 5):S79–S89.

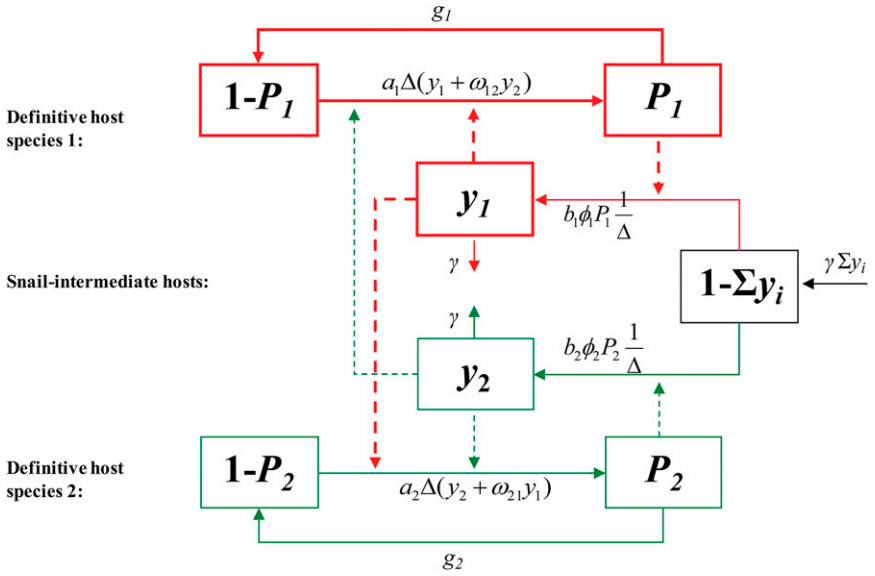

Fig. S1. Flow diagram of the multihost *S. japonicum* model. Parameter symbols are defined in Table S6. Only two definitive host species are shown for clarity. Solid arrows represent flows between compartments; dashed arrows represent transmission events between definitive and intermediate hosts.

Table S1. Census and parasitological survey data among definitive hosts for *S. japonicum* infection in Anhui, China

	Marshland region				Hilly region			
	GH	HP	XZ	Overall*	LQ	LS	YT	Overall*
Population sizes								
Humans	2,156	1,777	2,165	6,458	1,071	781	721	2,573
Cattle	292	27	51	370	1	7	0	8
Water buffalo	25	6	7	38	0	0	0	0
Goats	0	0	21	21	0	0	0	0
Dogs	20	18	36	74	43	29	17	89
Cats	30	3	1	34	25	45	43	113
Rodents (no. caught)	7	1	1	9	18	22	9	49
No. of trap-days	146	100	88	334	141	142	141	424
Rodents caught per trap-days	0.048	0.010	0.011	0.027	0.128	0.155	0.064	0.116
No. positive/no. examined								
Humans (IHA†)	325/1,303	107/788	91/1,084	523/3,175	107/553	76/493	62/343	245/1,389
Humans (stool examination)	1/325	0/75	1/88	2/488	0/96	5/51	3/52	8/199
Cattle	42/100	4/19	27/38	73/157	0/1	0/7	—	0/8
Water buffalo	3/25	0/6	1/7	4/38	—	—	—	—
Goats	—	—	11/20	11/20	—	—	—	—
Dogs	2/8	0/10	0/24	2/34	9/34	1/23	4/17	14/74
Cats	0/1	—	0/1	0/2 (1/67)‡	1/5	0/17	0/17	1/39
Rodents	0/7	0/1	0/1	0/9	6/18	5/22	2/9	13/49

GH, Guanghui; HP, Heping; LQ, Longquan; LS, Longshang; XZ, Xingzhuang; YT, Yuantou.
*Overall values derived from pooling data across all three villages in each region.
†IHA is the blood test for antibodies to *S. japonicum*.
‡Numbers in parentheses show data from cats in the marshland region from Wang et al. (1).

1. Wang TP, et al. (2005) Transmission of *Schistosoma japonicum* by humans and domestic animals in the Yangtze River valley, Anhui province, China. *Acta Trop* 96(2–3):198–204.

Table S2. Results of snail surveys of *O. hupensis* in Anhui, China

	Marshland region				Hilly region			
	GH	HP	XZ	Overall*	LQ	LS	YT	Overall*
Snail habitat area (m^2)	600,200	275,800	100,000	976,000	90,754	62,612	83,977	237,343
Snail density (no. per m^2)	6.7	6.8	10	7.1	25.7	16.2	22.5	23.9
No. examined for *S. japonicum* infection	1,222	632	355	2,209	7,683	1,823	3,349	12,855
No. positive for *S. japonicum* infection	6	14	2	22	17	10	24	51

GH, Guanghui; HP, Heping; LQ, Longquan; LS, Longshang; XZ, Xingzhuang; YT, Yuantou.
*Overall values derived from pooling data across all three villages in each region.

Table S3. Range and parameters for the $\beta(\alpha, \beta)$ prior distributions used for the sensitivities and specificities of diagnostic tests for *S. japonicum* infection in definitive and snail intermediate hosts

	Sensitivity			Specificity			
		β-coefficients			β-coefficients		
	Range (%)	α	β	Range (%)	α	β	Source
IHA test (humans only)	80–95	67.18	9.60	50–90	14.00	6.00	1–3
Hatching test	50–95	10.70	4.06	100*	NA	NA	1 and expert opinion
Snail crushing	85–95	71.25	3.75	100*	NA	NA	Expert opinion

NA, not applicable.
*Specificities of the miracidia hatching test and snail crushing technique were assumed to be known at 100%, and therefore, no prior distributions were necessary.

1. Yu JM, de Vlas SJ, Jiang QW, Gryseels B (2007) Comparison of the Kato-Katz technique, hatching test and indirect hemagglutination assay (IHA) for the diagnosis of *Schistosoma japonicum* infection in China. *Parasitol Int* 56(1):45–49.
2. Wang XH, Wu XH, Zhou XN (2006) Bayesian estimation of community prevalences of *Schistosoma japonicum* infection in China. *Int J Parasitol* 36(8):895–902.
3. Zhou YB, et al. (2007) Field comparison of immunodiagnostic and parasitological techniques for the detection of Schistosomiasis japonica in the People's Republic of China. *Am J Trop Med Hyg* 76(6):1138–1143.

Table S4. Distribution and parameters for the $\beta(\alpha, \beta)$ prior distributions used for the prevalence of *S. japonicum* infection in definitive and snail intermediate hosts in Anhui, China

	Mean	SD	β-distribution coefficients		Source
			α	β	
Humans	0.021	0.031	0.424	19.559	1 and NSD
Cattle	0.131	0.118	0.943	6.278	1 and NSD
Water buffalo	0.110	0.078	1.663	13.479	1 and NSD
Goats	0.106	0.128	0.511	4.306	1 and NSD
Pigs	0.012	0.030	0.152	12.195	1 and NSD
Dogs*	—	—	1	1	
Cats*	—	—	1	1	
Rodents*	—	—	1	1	
Snails	0.005	0.007	0.439	92.249	NSD

NSD, national surveillance data from 2005 collected in Anhui Province.
*Very little previous data were available for the prevalence of infection in dogs, cats, and rodents, and therefore, a vague (uninformative) prior was used.

1. Wang TP, et al. (2005) Transmission of *Schistosoma japonicum* by humans and domestic animals in the Yangtze River valley, Anhui province, China. *Acta Trop* 96(2–3):198–204.

Table S5. Median values (and 95% Bayesian credible intervals) for the adjusted prevalence (%) of *S. japonicum* infection among definitive hosts and snail intermediate hosts in Anhui, China

	Marshland region				Hilly region			
	GH	HP	XZ	Overall*	LQ	LS	YT	Overall*
Humans	0.1 (0.0, 0.6)	0.0 (0.0, 0.6)	0.2 (0.0, 0.7)	0.1 (0.0, 0.4)	0.0 (0.0, 0.7)	2.2 (0.8, 4.7)	1.5 (0.4, 4.0)	1.1 (0.5, 2.3)
Cattle	48.9 (36.4, 65.1)	22.2 (8.3, 44.9)	68.0 (51.6, 82.8)	54.5 (43.3, 69.5)	nd	5.6 (0.2, 28.3)	—	5.3 (0.1, 26.3)
Water buffalo	13.3 (4.6, 28.2)	7.1 (0.8, 24.4)	11.8 (2.5, 30.5)	12.6 (4.8, 25.7)	—	—	—	—
Goats	—	—	53.8 (32.0, 76.9)	53.8 (32.0, 76.9)	—	—	—	—
Pigs	0.0 (0.0, 7.2)	nd	0.0 (0.0, 5.9)	0.0 (0.0, 5.0)	nd	0.0 (0.0, 4.5)	0.0 (0.0, 4.9)	0.0 (0.0, 3.0)
Dogs	40.0 (10.1, 86.6)	9.0 (0.3, 42.8)	4.0 (0.1, 21.7)	10.8 (2.5, 29.6)	38.8 (19.8, 72.3)	9.8 (1.4, 32.4)	36.1 (13.3, 75.3)	27.7 (15.4, 51.2)
Cats	nd	nd	nd	3.5[†] (0.5, 12.3)	36.7 (5.9, 88.9)	5.3 (0.2, 27.8)	5.3 (0.2, 27.8)	6.0 (0.9, 20.2)
Rodents	11.7 (0.4, 55.1)	nd	nd	9.4 (0.3, 46.6)	48.4 (22.0, 87.1)	34.5 (14.1, 70.0)	36.2 (9.0, 82.6)	38.3 (21.5, 68.8)
Snails	0.5 (0.2, 1.0)	2.1 (1.1, 3.4)	0.5 (0.01, 1.5)	1.1 (0.7, 1.6)	0.2 (0.2, 0.3)	0.6 (0.3, 1.0)	0.8 (0.5, 1.1)	0.4 (0.3, 0.6)

Estimates are not given for cases where less than five individuals were examined (nd) or no individuals of that species were present (—). GH, Guanghui; HP, Heping; LQ, Longquan; LS, Longshang; XZ, Xingzhuang; YT, Yuantou.
*Overall values derived from pooling data across all three villages in each region.
[†]Prevalence among cats in the marshland region estimated using data from Wang et al. (1).

1. Wang TP, et al. (2005) Transmission of *Schistosoma japonicum* by humans and domestic animals in the Yangtze River valley, Anhui province, China. *Acta Trop* 96(2–3):198–204.

Table S6. Definitions and values of parameters for the multihost *S. japonicum* transmission model

Symbol	Description (units)	Distribution	Notes/references
Input parameters			
$1/g_i$	Duration of infection in species i (y)		
	Human	Tri (1, 0.5, 1.5)	Annual rate of chemotherapy in humans
	Water buffalo	Unif (2, 5)	3–5 y average lifespan of adult parasites (1)
	Cattle	Unif (2, 5)	1–2 y average lifespan of rodents (2, 3)
	Goat	Unif (2, 5)	
	Cat	Unif (2, 5)	
	Dog	Unif (2, 5)	
	Rodent	Unif (1, 2)	
$1/\gamma$	Infected snail lifespan (y)	Unif (0.082, 0.433)	30–158 d average lifespan of infected snails (1, 4, 5)
ε_i	Mean intensity of infection of infected host of species i (eggs g^{-1} feces)		
	Human	Unif (0.1, 24)	
	Water buffalo	Unif (0.1, 10)	Local data (6, 7)
	Cattle	Unif (0.1, 10)	
	Goat	Unif (0.5, 25)	
	Dog	Unif (0.5, 20)	
	Cat	Unif (0.1, 10)	
	Rodent	Unif (100, 600)	
θ_i	Rate of fecal excretion of definitive host of species i (g d^{-1})		
	Human	Tri (160, 100, 325)	
	Water buffalo	Tri (14,667, 9,800, 21,550)	6, 8
	Cattle	Tri (5,967, 4,450, 7,650)	
	Goat	Tri (191, 95, 280)	
	Dog	Tri (99, 45, 150)	
	Cat	Tri (20, 7, 53)	
	Rodent	Unif (1, 2)	
ω_{ij}	Level of mixing between definitive host species i and j relative to mixing within host species i	$\begin{cases} 1 & \text{if } i=j \\ \text{Unif}(0.01,\ 1) & \text{otherwise} \end{cases}$	Used to incorporate spatial structuring of definitive host species (set wide)
φ_i	Density of definitive hosts of species i (no. individuals m^{-2} snail habitat)	Data estimate	Local data (Table S1)
Δ	Snail density (no. snails m^{-2} snail habitat)	Data estimate	Local data (Table S2)
$Y*$	Steady state prevalence in snails	Posterior distribution from Bayesian analysis of local data	Local data (Table S5)
P_i^*	Steady state prevalence in definitive host species i	Posterior distribution from Bayesian analysis of local data	Local data (Table S5)
Parameters estimated from model			
a_i	Rate of incidence for a single definitive host of species i per unit density of snails infected by host species i (y^{-1})	Estimated from model	$a_i = \frac{g_i P_i^*}{\Delta(1-P_i^*)\sum_j \omega_{ij} Y_j^*}$
b_i	Rate at which an infected definitive host of species i causes snail infections (y^{-1})	Estimated from model	$b_i = \varepsilon_i \theta_i \eta$
η	Probability that an egg excreted by an infected definitive host causes a snail infection	Estimated from model	$\eta = \frac{\Delta\gamma Y^*}{\sum_i (\varepsilon_i \theta_i \varphi_i P_i^*)(1-Y^*)}$

Tri, triangle distribution (mode, minimum, maximum); Unif, uniform distribution (minimum, maximum).

1. Anderson RM, May RM (1991) *Infectious Diseases of Humans: Dynamics and Control* (Oxford Univ Press, Oxford).
2. Jackson WB (1982) Norway rat and allies. *Wild Mammals of North America: Biology, Management, Economics*, eds Chapman JA, Feldhamer GA (Johns Hopkins University Press, Baltimore).
3. Nowak R, Paradiso J (1983) *Walker's Mammals of the World* (Johns Hopkins University Press, Baltimore, MD), 4th Ed.
4. Sun LP, et al. (2004) [Effect of *Schistosoma japonicum* infection on *Oncomelania hupensis*.]. *Zhongguo Xue Xi Chong Bing Fang Zhi Za Zhi* 16(4):265–268. Chinese.
5. Liang S, Spear RC, Seto E, Hubbard A, Qiu D (2005) A multi-group model of *Schistosoma japonicum* transmission dynamics and control: Model calibration and control prediction. *Trop Med Int Health* 10(3):263–278.
6. Wang TP, et al. (2005) Transmission of *Schistosoma japonicum* by humans and domestic animals in the Yangtze River valley, Anhui province, China. *Acta Trop* 96(2–3):198–204.
7. Lu DB, et al. (2010) Contrasting reservoirs for Schistosoma japonicum between marshland and hilly regions in Anhui, China—a two-year longitudinal parasitological survey. *Parasitology* 137(1):99–110.
8. Ebihara K, Shiraishi R, Okuma K (1998) Hydroxypropyl-modified potato starch increases fecal bile acid excretion in rats. *J Nutr* 128(5):848–854.