Supporting Information

Materials and Methods

Study area

Iquitos is the principal city in the department of Loreto located in Northeast Peru. The nearest cities with significant dengue transmission are Yurimaguas and Pucallpa, both ~ 2 days away by river. Numerous, smaller communities in the area are connected by road or river. The climate is moderately seasonal, with reduced precipitation between June and August and slightly warmer temperatures between September and February. Iquitos has been the site of dengue research since the early 1990‘s with continuous community-based, cohort research by our team since 1999 (1-12). Dengue transmission typically occurs between October and April of each year (12). In 1990 DENV-1 invaded Iquitos (1), followed by DENV-2 (American) in 1995 (3), DENV-3 in 2001 (11), and DENV-4 in 2008 (13). In general, dynamics have been dominated by single serotypes (12). The LRHD attempts to control vector populations by treating water-holding containers with larvicide, removing potential development sites from households, and through education campaigns. In the event of an outbreak, the LRHD conducts city-wide fumigation campaigns using non-residual pyrethroid insecticides.

Longitudinal cohort

Based on differences in historical records of DENV transmission and relative geographic independence (local availability of services such as schools and health clinics), two previously described neighborhoods, Maynas (MY) and Tupac Amaru (TA; 11) were selected for febrile surveillance and longitudinal monitoring of anti-dengue virus neutralizing antibodies (Fig. 1A). These neighborhoods exhibit modest differences in socio-economic status and population density. In TA, 60% of houses were constructed of more rustic materials (e.g., wood) versus 53% in MY. Also, the median number of household residents was 6 in TA versus 5 in MY (not statistically different, $p = 0.06$; neg. binomial: mu = 6.16 residents per house, theta = 9.07). Enrollment in the study was offered to people living in all households on contiguous blocks with an estimated population of ~3000 in each neighborhood between November 2007 and May 2008 until we recruited 2,444 longitudinal participants ≥5-years old, divided evenly between the two neighborhoods. Over the study period we took blood samples from longitudinal participants (≥5 y old) at ~ 6 month intervals to test for serological evidence of DENV infection (Fig. S1; Table S3). Baseline samples were taken before cluster investigations started in August 2008.

Febrile surveillance

In April 2008, active community-based surveillance was initiated in all consenting households based on a strategy described previously (8), where study personnel visited each household a minimum of 3 days per week to ask if anyone living there had a febrile illness. When a febrile person (> 3 years old) was identified, written informed consent was obtained from adult participants or the parents of participants <18 years old. Assent was obtained from participating children (8-17 years old). Blood was drawn at the time of case identification (acute sample), followed by a convalescent blood sample 14-21 days later. We offered 2 options for study participation: 1) acute and convalescent blood samples with clinical evaluation by a project physician and 2) option #1 procedures with an interview inquiring about the participant’s movement patterns over the previous 15 days (Fig. S1). In the latter situation, we asked for permission to visit houses identified in the interview. These individuals thus served as the 'index cases' of contact-site cluster investigations.

Contact-site cluster investigations

Febrile individuals captured by our community-based surveillance who agreed to option #2 of our consent process (‘index cases’; Fig. S1) were interviewed by trained field technicians at the earliest, convenient time for the participant and family. Usually (92% of the time) interviews were conducted the same day the case was detected or on the subsequent day. Interviews were developed in Iquitos to be culturally sensitive and

effective at identifying locations visited in the previous 15 days and were evaluated in comparison to GPS movement monitoring in a separate component of the study (Paz-Soldan in preparation; 9, 10). A 15-day period was chosen because, in most cases, it would include the time of infection. Intrinsic incubation of DENV is 4-7 days before viremia develops and symptoms begin approximately a day after viremia onset (14, 15). We commonly detected febrile people on day 2 or 3 of symptoms, which are generally thought to last ~5 days. On average we expect participants were infected 7 to 11 days previously (15). The movement interview was structured to identify commonly visited, non-residential locations and households. Non-residential locations included parks, plazas, markets, churches, schools, and other such places where people aggregate, but do not live. Households were homes of friends and relatives or mixed-use locations, such as small stores, restaurants and bingo parlors where people both work and live. For logistical and practical reasons, including documented low mosquito abundances in non-residential places (7), we focused our cluster investigations on households. After completion of the movement interview, our team identified and geo-coded the contact locations described by the index case. Sites outside of Iquitos (4% of all places identified in interviews) were excluded for logistical reasons. Residents of contact locations falling within the urban area of Iquitos and ≥5 years old were then informed about the study and offered the opportunity to participate by providing 2 blood samples ~15 days apart. In the second season of the study, contacts residing within the neighborhoods were additionally invited to participate in a movement interview. We revisited contact houses every other day over the 15 days of the cluster investigation to follow up with participants and inquire about people with new symptoms. Contacts developing fever during a cluster investigation were offered the opportunity to participate as febrile cases (option #1 of our consent process). Consenting individuals provided a blood sample at that moment and a convalescent sample ~15 days later. In these cases we did not take a convalescent blood draw 15 days after the acute cluster sample, so individuals provided a maximum of 3 samples as part of a cluster investigation. If cluster investigations coincided with our scheduled longitudinal blood draws, we would use the acute cluster sample of cohort participants for acute serology and PRNT to limit the number of samples requested.

Cluster initiation

Our aim was to balance clusters initiated by DENV+ (cases) and DENV- (controls) febrile people within 2 weeks in the same neighborhood. This depended, however, on local DENV transmission, participant consent, timing of samples and results, and programmatic logistics. The homes of index cases captured within a 2-week period were restricted to a minimum separation of 50 meters to avoid geographic overlap and limit spatial correlation due to mosquito movement. In the first season we had fewer DENV- control clusters than DENV+, in part due to an elevated DENV-4 incidence rate in our study neighborhoods in October of 2008. Subsequently, the LRHD conducted a city-wide intervention (indoor ultra-low volume application of pyrethroids) in the last week of October (Fig. 1C). The affected area included ~140,000 residents (~33,000 houses; data provided by LRHD). After spraying, transmission was reduced in our study areas and then further reduced by a second intervention in February 2009 (~42,000 houses). In the second season, we better balanced DENV+ and DENV- clusters, which was helped by overall reduced transmission rates, up until the Ministry again performed a city-wide intervention during March 2010 (~42,000 houses; Fig. 1C).

For a DENV+ cluster, the index case was positive by RT-PCR or IFA in the acute sample or showed a 4-fold rise in IgM titer between acute and convalescent samples. Otherwise the cluster was classified as DENV-. In several instances, we could not classify a cluster as positive or negative due to uncertainty in the index case's serology. In one instance we did not get a control sample for IgM. In 2 cases IgM was elevated in the acute sample during the last quarter of 2008, which could indicate exposure up to several months before and to DENV-3. In 3 cases $PRNT_{70}$ results for the acute cluster sample indicated seroconversion since the prior longitudinal sample, but acute serology was negative. Thus we do not know when seroconversion occurred in the >7 month interval between longitudinal blood draws in these cases.

Sample processing

Longitudinal blood samples were analyzed by plaque reduction neutralization tests with a 70% cutoff for all 4 DENV serotypes as generally described elsewhere ($PRNT_{70}$; (11)). Here, however, all samples were diluted 4-fold prior to use: 1:40, 1:80, 1:120, and 1:640. Acute samples from febrile cases and cluster participants were inoculated onto C6/36 cells for virus isolation as previously described (12). Acute and convalescent samples were assayed for DENV-specific IgM antibodies by antibody-capture ELISA (12). Individuals opting for option #2 also had their acute blood tested for DENV by RT-PCR (12) within 48 hours of the blood draw in order to initiate cluster investigations.

Classification of serology results

The final result for individual cluster participants was determined by considering outcomes of all laboratory assays and classifying a person has having a recent infection, acute infection, prior exposure, or no prior exposure (serologically negative). All cluster participants whose serum was positive by PCR or IFA in either the acute or convalescent sample were classified as acute infections. Participants evidencing a 4x rise in IgM titer between the acute and convalescent were also considered acute infections. If the acute sample was taken a few days after defervescence, a 4x rise in IgM titer in the second sample could have been observed. This, plus the period of viremia (2-14 days) and intrinsic incubation (4-7 days) indicates we could detect acute infections that occurred ~20 days before the start of the cluster. For this reason and because they could be connected to infections in clusters, febrile cases we detected in cluster households within 30 days of study initiation who were positive for DENV by PCR or IFA or exhibited IgM seroconversion were included as acute infections. If IgM was elevated in the acute sample without a rise in the second sample (and no other evidence the infection occurred within 30 days of the cluster) the participant was classified as a recent infection. Evidence for prior exposure to DENV-4 was based on detection of neutralizing antibodies by serotype-specific PRNT of serum from individuals participating in the longitudinal cohort. If neutralizing titers were elevated above a minimum cutoff in a sample taken prior to the cluster start date, the individual was classified as previously exposed. The serotype-specific cutoffs were 1:60 for DENV-1, 1:80 for DENV-2, 1:60 for DENV-3, and 1:40 for DENV-4. Taking results for all serotypes in aggregate, assuming predominant DENV-4 transmission based on febrile surveillance data and using serum samples taken before the invasion of DENV-4, we estimated 85% sensitivity and 99% specificity of our ability to identify DENV-4 seroconversions by $PRNT_{70}$. Our estimates of seroprevalence, and thus susceptibility to DENV-4, that are based on PRNT are thus conservative. Cluster participants who had been identified as DENV-4 positive febrile cases at a time before (>30 days) the cluster was initiated, but for whom we did not have serological data, were classified as previously exposed at the time of the cluster. Remaining individuals for whom we had a sample were considered negative for DENV-4 and thus susceptible.

Data analysis

Estimating individual risk

Ae. aegypti is day-active and opportunistic (16, 17), so all locations visited potentially contribute to an individual's risk of infection if mosquitoes are present (18). The probability of a successful encounter with the pathogen (effective transmission of virus from mosquito to human) will thus increase with exposure across an individual's activity space (the sum of locations visited recently). We approximated exposure to DENV for index cases by counting the number of concurrent acute infections in households. Although we expect individuals to be exposed to virus anywhere in their activity space, infection is still very likely to occur in the home. Thus, the number of infections observed in a house on the day when the index case is detected will correlate with the number of infective mosquitoes that were in the house 7 - 20 days prior when the infections took place.

We first related risk to the product of the number of acute infections in houses, I_j, and time spent at the house:

$$\pi_i \propto \sum_j^J I_j \cdot f_{i,j} \cdot t_{i,j}$$

where $f_{i,j}$ is the number of visits by person i to house j and $t_{i,j}$ is the time spent during visits. We were limited to only houses outside the home, however, because we did not collect time in house information during the first season. We do not know what the actual function for risk in sites should be. Although *Ae. aegypti* is thought to respond quickly to the appearance of a host, suggesting greater risk with more frequent visits, it may be that during very short visits people remain active and thus are difficult for mosquitoes to bite. For this reason we evaluated this simple relationship and another excluding $t_{i,j}$ with a non-parametric statistic (Wilcoxon, 1-tailed test) to test for a difference in exposure outside the house between DENV+ and DENV- index cases.

Because we were limited in our ability to estimate contact with mosquitoes in terms of the actual time spent at houses in the activity space, we developed an index that is a form of the Shannon-Wiener diversity index (see Eq. 1 in main text;19). The estimation of σ_i included all acute infections, including the index case. Because this could be considered a bias, however (i.e., because exposure for DENV+ clusters will always be higher than DENV- clusters), we also estimated σ_i excluding the index case (Table S4). Any non-zero value could be used for the correction factor c, which determines the slope of the relationship between J and σ_i. We set c to $(J_{max} - 1) / (2 \bullet \ln (J_{max}))$ to make the distribution of values for σ_i symmetric with range $-\ln (J_{max})$ to $\ln (J_{max})$. For the data presented here with $J_{max} = 9$, σ_i could range from -2.197 to 2.197.

Models

We estimated risk as $logit\ (\pi_i) = \alpha + \beta_1 \bullet \sigma_i + ... + \beta_n X_{n,i} + \varepsilon$, and fitted models with alternative parameter combinations to the data using maximum likelihood (Tables S3 and S4; Generalized linear models [GLM] with binomial error distribution and logit link function). No random effects were included in cluster level models.

Attack rates were calculated as the number of acute infections over the number of susceptible individuals: acute / (acute + negative). We did not include recent infections unless otherwise indicated, because of the ambiguities associated with elevated IgM titer in the first sample. During the first season, elevated IgM could have been caused by DENV-3 infection because IgM titers remain elevated for long periods, up to 5 months in our case. Recent infection with one serotype may provide a brief period of cross-protection against other serotypes, however, and so we considered all recent infections to be effectively immune to DENV-4 at the time of the cluster. Seroprevalence rates were estimated as the proportion of participants providing bi-annual blood samples with DENV-4 neutralizing antibody detected by $PRNT_{70}$ assays on samples taken prior to the cluster and without evidence (IgM) of recent DENV-4 infection.

Attack and seroprevalence rates and their deviance due to both fixed and random effects were estimated and examined using GLM, General Linear Mixed-effect Models (GLMM), or General Additive Models (GAM) and General Additive Mixed-effects Models (GAMM) where continuous covariates were included (Tables S3, S5-10). All models assumed binomial error distributions and logit link functions and were fit by Maximum Likelihood. Over-dispersed data were fit with a binomial model with dispersion parameter, $\hat{c} > 1$ ('quasibinomial'). The code of individual cluster investigation was included as a random effect (on the intercept) in household-scale models, but was not retained because it did not improve model fit. Estimation of house-level attack rates also included an offset term indicating whether the house participated in the longitudinal cohort, where we could estimate the proportion of residents susceptible to DENV-4 at the beginning of the cluster investigation.

The proportion of inapparent infections was estimated per cluster for participants living in houses under active, community-based surveillance with a binomial GLM including cluster type as the only effect (Table S2).

For this estimate we are primarily interested in DENV+ clusters, because the result can be interpreted as the proportion of new infections with symptoms associated with an apparent infection. We included DENV- clusters for comparison, but note that there is an inherent bias in the estimation of inapparent rates in DENV+ versus DENV- clusters that is due to the differences in transmission rates. Transmission rates are very low in DENV- clusters, introducing much higher sampling variation. The true rate of inapparent infection should fall somewhere between that estimated for DENV+ and DENV- clusters.

Seasonally forced transmission model

We developed a seasonally forced transmission model to estimate the additional risk of infection due to association with the activity space of a DENV infected person. Briefly, we partition infection risk into 2 components: 1) a 'within-cluster' risk derived from residing in the activity space of a DENV infected person and 2) a background risk associated with living in Iquitos (estimated from negative clusters). This model was then fit to the data from contact-site cluster investigations using Likelihood and Bayesian approaches (see SI for additional details).

Model selection

All possible parameter combinations were not considered because of sample sizes. We a priori postulated multiple competing and semi-nested models based on our hypothesis and prior information on potentially important factors influencing transmission. In all cases models were evaluated with Akaike's Information Criterion (AIC) with correction for small sample sizes (AICc and QAICc for cases of over-dispersion) and, where nested, with log-likelihood ratio tests (Chi-square tests for binomial models, F tests for 'quasi'-binomial models; 20). Generally the best model based on Akaike weights, w_i, was used, although models with very small differences in AIC_c (e.g., $\Delta AIC_c < 2$) are arguably equally good fits to the data (21). The importance of individual effects was evaluated with odds ratios or predicting rates with estimates of standard error (s.e.m). All models were evaluated by assessing the distribution of residuals relative to predicted values and as a function of all main effects.

All models and diagnostics were performed in R, using the mgcv, lmer, MASS, and spatstat packages (22-26).

Transmission model formulation

Model Assumptions

The model is based on a few assumptions about the dynamics in DENV+ and DENV- clusters:

1. Due to the fact that *a priori* it must be true that the index individual for a DENV+ cluster is infected and the index individual for a DENV- cluster is not infected, we remove them from all calculations of incidence within the cluster. Were we to not remove them from these calculations, we would bias our DENV+ cluster to have more infections than our DENV- clutsters (specifically 1 more infection).

2. For each cluster k, from the remaining N_k members of the cluster who potentially may be susceptible (i.e. for whom we do not have knowledge that they have already had dengue), each person's past infection status is independent of each other person's status and the probability of each person being susceptible, $p_k^{(1)}$, is constant within the cluster. This makes the number of susceptible people in a cluster, S_k, binomially distributed
$$S_k \sim \text{Binom}\left(N_k, p_k^{(1)}\right).$$

3. For each DENV- cluster k, if there are S_k susceptible members, each persons current infection status is independent of each others, and the probability of each being infected, $p_k^{(2)}$, is constant within the cluster. This probability essentially captures the chance of a randomly selected person being infected (or without cluster infection). This makes the number of infected people in a DENV- cluster with S_k susceptible members, $I_k^{(-)}$, binomially distributed
$$I_k^{(-)} \sim \text{Binom}\left(S_k, p_k^{(2)}\right).$$

4. For each DENV+ cluster k, infections can occur two ways: first through 'without cluster infection' governed by $p_k^{(2)}$, and second through within cluster infection. We assume that each members can only be infected through one method (within or without cluster), each susceptible members current infection status is independent of each other person's status and the probability of each being infected from within their own cluster, $p_k^{(3)}$ is constant within the cluster.

Both assumption 2 and 3 assume that infections within a cluster are independent of other infections within the same cluster which is somewhat unrealistic since it is possible that members of a cluster may move throughout the city more frequently that two randomly chosen people. Assumption 2 implies that no infection that occurs in a DENV- cluster came from an earlier infection within the cluster, which is also somewhat unrealistic since it is possible to have started a DENV- cluster with a negative member of a cluster that is actively undergoing a cluster-wide epidemic. Assumption 1 and 2 imply that, for a DENV- cluster k with N_k members whose past infection status is unknown, the number of active infections, $I_k^{(-)}$, is binomially distributed
$$I_k^{(-)} \sim \text{Binom}\left(N_i, p_k^{(1)} \cdot p_k^{(2)}\right).$$

Assumption 1, 2 and 3 imply that, for a DENV+ cluster k with N_k members whose past infection status is unknown, the number of active infections, $I_k^{(+)}$, is binomially distributed
$$I_k^{(+)} \sim \text{Binom}\left(N_i, p_k^{(1)} \cdot \left(p_k^{(2)} + p_k^{(3)} - p_k^{(2)} \cdot p_k^{(3)}\right)\right).$$

Relating epidemiological and model dynamics

To identify relationships between dengue dynamics within clusters, both between districts and from year to year, we allow the three probabilities that govern the dynamics to be functions of space and time. To link the model to the outbreak dynamics, we relate the first two probabilities to the proportion of infected and susceptible people in the city through time. Specifically, assuming that the entire population was susceptible August 1st, 2008 (what we call $t = 1$), we let the percent of the population that is susceptible, $s(t)$, and actively infected, $i(t)$, be defined as follows (for $1 \leq t \leq 720$):

$$i(t) = f(t) \tag{S1}$$

$$s(t) = 1 - \int_0^t i(t')dt' \tag{S2}$$

where $f(t)$ will be selected through model selection from a variety of candidates below (note that we could write the above equations as a system of ODEs, but for this analysis, nothing is gained by doing so). Then we define, for cluster i that was sampled at time t_i, $p_i^{(1)}$ and $p_i^{(2)}$ as

$$p_i^{(1)} = s\left(t_i\right) \tag{S3}$$

$$p_i^{(2)} = \alpha \cdot i\left(t_i\right) \tag{S4}$$

To allow these probabilities to be functions of space as well, we let f and α be different for each district. Specifically, we detail the various spatio-temporal relationships we test in the following subsection.

Candidate models

To model the seasonality of dengue throughout the city, we tested using sinusoidal function of period 1 year and/or 6 months. To allow infections to occur throughout the year, we force percent infected to be always greater than zero positive. A relatively basic example of a choice for f (for the two years we are interested) is

$$f(t) = \begin{cases} C + \beta_1 + \beta_1 \cdot \cos\left(\pi \frac{t-150}{180}\right), & 1 \leq t \leq 330; \\ C + \beta_2 + \beta_2 \cdot \cos\left(\pi \frac{t-150}{180}\right), & 330 \leq t \leq 720; \end{cases} \tag{S5}$$

This f has a different outbreak level the first year than the second. To allow for more flexibility, we also consider a linear combination of a sinusoidal function of period 1 year and one of period 6 months (to get more flexibility in the functions we can approximate using essentially wavelet analysis). For brevity, if we just used the 6 month periodic seasonality, the function f would be

$$f(t) = \begin{cases} C + \beta_1 + \beta_1 \cdot \cos\left(\pi \frac{t-60}{90}\right), & 1 \leq t \leq 150; \\ C + \beta_2 + \beta_2 \cdot \cos\left(\pi \frac{t-60}{90}\right), & 180 \leq t \leq 330; \\ C + \beta_3 + \beta_3 \cdot \cos\left(\pi \frac{t-60}{90}\right), & 330 \leq t \leq 510; \\ C + \beta_4 + \beta_4 \cdot \cos\left(\pi \frac{t-60}{90}\right), & 510 \leq t \leq 720; \end{cases} \tag{S6}$$

When we use both periods, we have up to 7 parameters. As mentioned above, we can test spatial effects by allowing these parameters to be different for the two different districts (though then the seasonality would be governed by 14 parameters).

We can also consider temporal and spatial effects on the chance of getting a within cluster infection by allowing $p^{(3)}$ to be different for both space and time. Due to the amount of data, we chose only to test $p^{(3)}$ as a different constant within each district and within time, but there is no reason that it couldn't be a dynamic function of time as well.

Results and Discussion

Summary of clusters

Results from contact-site cluster investigations initiated by index cases with indeterminate diagnosis are shown in Tables S1 and S2 (classified as 'ND') and in descriptions of individual movement patterns, but otherwise are excluded from statistical analyses comparing exposure and transmission rates (for n = 48 DENV+ and DENV- clusters in models). We excluded 17 other clusters from this analysis (9 from the first season, 8 from the second): 5 clusters initiated by DENV-3 infected cases (n = 5), 1 cluster without a DENV diagnosis for the index case, 1 cluster that overlapped with another, and 10 clusters with no participation outside of the index home. Of these 10, in 6 instances no contacts living in contact-sites participated and in the other 4 instances no additional households were reported as visited recently, thus we could not assess the extent of exposure outside of the home of the index case. Based on a much larger interview data set (n = 481), 85.3% of febrile participants report having visited at least 1 additional house in the previous 15 days (mean = 2 houses). We concluded that the clusters we analyzed were representative of movement by the majority of people in Iquitos.

Participants

For the period of study, twice as many index cases were residents of the TA neighborhood (n = 36; Fig. 1A, 1B). Index cases included adults (n = 22) and children (<18 years old; n = 32) and we had a higher percentage of females (63%) and children (59%), though these were not statistically significant (χ^2=0.007, p=0.93). There was no systematic variation in the number of contacts nor participation rates by type of contact cluster (DENV+ or DENV-), neighborhood, or season (Poisson and binomial GLM). Among contacts, there was under-representation of adult males relative to the other groups (χ^2=5.731, p = 0.016; Fig. S2). Overall, 54.6% of contact households were participating in our long-term, prospective dengue cohort and 43.8% of study participants had provided bi-annual blood samples for analysis by PRNT to detect the presence of pre-existing, neutralizing antibody for DENV-4.

Serological results

A summary of serological results is in Table S1. Most contacts identified with acute DENV infections were likely exposed around the same time as the index case because they were identified by either virus isolation in the acute blood draw (37%) or by IgM seroconversion (50%). For households falling within neighborhoods under febrile surveillance, we estimate that 17.9% of infections among contacts were apparent (body temperature >38ºC; Table S3).

Exposure

Total time spent at contact houses (frequency * duration; not including the home of the index case) did not differ between DENV+ and DENV- index cases (median = 1410 minutes, range = 0 - 14820 minutes). Exposure at contact houses, taken as the product of the number of acute infections in a house and total time spent there, tended to be higher in DENV+ clusters (median log(exposure) = 3.37 vs. 2.71; p = 0.053, Wilcoxon Test). Dropping time in the house and using only number of visits provided a slightly better result (p = 0.045). We were limited in this analysis, however, because it was common to have no acute infections in houses and because we did not have complete information on time in the home. Also, it is unclear that total time at a location is a good correlate of exposure. In seasons subsequent to those analyzed here we gathered more detailed information that will be used in the future to more carefully explore the actual relationship between visitation and exposure.

To illustrate the properties of the index of contact, σ, we fit it to the number of total acute infections in an activity space and the proportion of houses with ≥1 acute infection (infestation) with a GAM (Fig. 2B). The best model (adjusted R^2 = 76.7%) included a thin plate regression spline smoothing function of total acute infections (df = 2.1; p = 0.0001) and a linear fit to infestation (β = 0.8, p = 0.01). The index thus increases with more

transmission and more spread of transmission across an activity space. σ was greater for male children (<18 years old) relative to female adults ($p = 0.045$) and female children ($p = 0.06$), but not to adult males (p=0.44; 2-way ANOVA and Tukey's Honestly Significant Difference), suggesting elevated risk for young men relative to females. Despite the very strong effect of σ in the risk model, there were important outliers (Fig. S3). In several instances, we detected no other acute infections in the activity space of DENV+ individuals (n = 6), though 4 of these included individuals recently infected with DENV. We also observed DENV activity in DENV- clusters. Three DENV- clusters were initiated by index cases immune to DENV-4 at the time of detection. In one, recent and acute attack rates were elevated, suggesting the index case may have seroconverted recently.

Recent infection attack rates

In addition to the acute attack rates presented in the main text, we combined acute and recent (elevated IgM in first sample) infections in a 'recent' attack rate. This produced the same general result, with higher rates in DENV+ clusters (Tables S8 and S13). We recognize, however, that there is considerable uncertainty in estimating the actual timing of infection based on IgM in the absence of virus isolation data (see Methods).

Transmission model

For both model selection and parameter estimation we used both frequentist and Bayesian approaches. For the frequentist approach, we evaluated the maximum likelihood estimate for each model, compared models using either likelihood ratio tests (when one model was a sub-model of another model) or AIC (when neither model was a special case of the other). Once we arrived at a final model, we evaluated uncertainty through the use of profile likelihood based confidence intervals (note: these CI's hold asymptotically, so they are conservative here). For the Bayesian approach, we use the Metropolis-Hastings MCMC algorithm for parameter estimation and use DIC (deviance information criterion) to compare models. Then we use the posterior distribution of the final model to compute credible intervals (with the mean in the center of the CI).

The 1 year periodic seasonality model was greatly outperformed by a model that only had a 6 month periodic seasonality. This could be due in part to the low number of cases in DENV- clusters in the first half of the 2009-2010 season (i.e. Aug-Dec 2009). As such the final model is not the linear combination of the sinusoidal functions, and we only use the seasonality function described in Eq. (S6).

Referring to Eq. (S6), we found that there are significant differences between the first and second half of the 2008-2009 season for both neighborhoods (i.e. β_1 and β_2 are different for TA and MY, both with respect to each other and across space). For the second season there was also a difference between the two halves of the season, but this difference was not significantly different between the two neighborhoods (i.e. β_3 and β_4 were significantly different, but not statistically different between TA and MY; Table S14). As mentioned above, for identifiability, the within-cluster infection rate $p^{(3)}$ was assumed constant throughout a season, and could only possibly have been different between seasons and neighborhoods. We found that there is no significant difference in the within-cluster infection rate between seasons for TA and the within-cluster infection rate for the first season of MY was statistically indistinguishable from the TA rate (Fig. 4C). On the other hand, the within-cluster infection rate for the second season of MY was estimated to be zero (due to the fact that no DENV+ cluster over this period in MY contained any other positive cases other than the index case). Model behavior, predictions, and parameter estimates are shown in Figures 4C, 4D, Fig. S10 and Table S14. Fig. S9 shows the resulting percent susceptibility calculated using Eq. (S2).

Once we accounted for the various spatial and temporal heterogeneities within the data, we conducted a residual analysis to investigate the effect of distance. When we looked at the residuals (observed cases - predicted cases) and plotted these values against average distance between members of the cluster and the index person's home (Fig. S6), we found that they are essentially uncorrelated ($r = -0.0961$). This suggests that, while our model does not capture all of the dynamics that govern dengue transmission within and without clusters, the

distance contacts are from the house of the index case does not play an important role in explaining any more of the dynamics.

Acknowledgements

We first thank the residents of Iquitos for welcoming us into their homes and permitting these studies to be conducted. We greatly appreciate the support of the Loreto Regional Health Department including Hugo Rodríguez Ferruchi, Carmen Montelvan, Cristiam Carey, Carlos Manrique de Lara Estrada, Carlos Coral, Yuri Alegre, and Ernesto Curto who all facilitate our work in Iquitos. Dr. Moises Sihuincha at the Hospital Apoyo de Iquitos was instrumental in initiating these studies. We thank our Peruvian field team, Helvio Astete Vega, Wilder Carrasco Huamán, Esther Jennifer Ríos López, Shirley Maribel Guédez Gonzales, Wendy Lorena Quiroz Flores, Diana Bazaán Ferrando, Gabriela Vásquez La Torre, Isabel Bazán Arista, Rosa Regina Fernandez Montano, Rebeca Carrión, Wieslawa Álava Flores, Llerme Armas Pisco, Sadith Jovita Ricopa Manuyama, María Edith Juárez Baldera, Acela Rosario Mafaldo García, Jenny Fílida Gonzáles Sangama, Rosa Tamani Babilonia, Karina Chuquipiondo Vásquez, Rina Gonzales Jaba, Yolanda Torres Arévalo, Naida Rocío del Río Chávez, Luz Angélica Galvez Huayllahua, Leslye Angulo, Patricia del Carmen Barrera Bardales, Lupe Flore, Alex Vásquez, Claudio Rocha, Jorge Vásquez Belchoir, Alan Lozano, John Ramirez, Angelo Mitidieri, Rommel Vásquez Álvez, Jimmy Espinoza Benevides, Jimmy Maykol Castillo Pizango, Rusbel Huiñapi Tamani, Juan Luiz Sifuentes Rios, Néstor José Nonato Lancha, Federico Reátegui Viena, Victor Eléspuru Hidalgo, Edson Pilco Mermao, Abner Enrique Varzallo Lachi, Fernando Chota Ruíz, Ángel Puertas Lozano, Guillermo Iñapi Huamán, Fernando Espinoza Benevides and Manuel Ruiz Rioja for their enthusiastic and dedicated assistance in our efforts to study the movements of residents of Iquitos and the epidemiology of dengue. Dr. Greg Martin, Dr. John Sanders, and Roxana Lescano of the U.S. Naval Medical Research Center in Lima, Peru, were instrumental in facilitating these studies. Carolina Guevara, Zonia Ríos, and Angélica Espinoza provided laboratory support in Lima. We additionally thank Leslie Sandberg and Yui Yin Chu for administrative support. Kelly Liebman, Kanya Long, and Jacklyn Wong provided valuable feedback over the course of the study. William Reisen and Bborie Park at the Center for Vectorborne Diseases, UC Davis, provided technical resources and support. Chris Dye, Charles Godfray, Simon Hay, David Smith, and Helen Wearing kindly commented on the manuscript. Development of the ideas presented here benefited from the Training Workshops on the Ecology and Evolution of Infectious Diseases, funded in part by the National Science Foundation (EF 0722115), and the Research and Policy for Infectious Disease Dynamics (RAPIDD) program of the Science and Technology Directorate, U.S. Department of Homeland Security, and the Fogarty International Center, National Institutes of Health.

SI Appendix References

1. Phillips I et al. (1992) First documented outbreak of dengue in the Peruvian Amazon region. *Bulletin of the PanAmerican Health Organization* 26.
2. Hayes CG et al. (1996) The epidemiology of dengue virus infection among urban, jungle, and rural populations in the Amazon region of Peru. *American Journal of Tropical Medicine and Hygiene* 55.
3. Watts DM et al. (1999) Failure of secondary infection with American genotype dengue 2 to cause dengue haemorrhagic fever. *Lancet* 354.
4. Getis A, Morrison AC, Gray K, and Scott TW (2003) Characteristics of the spatial pattern of the dengue vector, Aedes aegypti, in Iquitos, Peru. *American Journal of Tropical Medicine and Hygiene* 69.

5. Morrison AC et al. (2004) Temporal and geographic patterns of *Aedes aegypti* (Diptera: Culicidae) production in Iquitos, Peru. *Journal of Medical Entomology* 41:1123-1142.
6. Schneider JR, Morrison AC, Astete H, Scott TW, and Wilson ML (2004) Adult Size and Distribution of Aedes aegypti (Diptera: Culicidae) Associated with Larval Habitats in Iquitos, Peru. *Journal of Medical Entomology* 41:634-642.
7. Morrison AC et al. (2006) *Aedes aegypti* (Diptera: Culicidae) production from non-residential sites in the Amazonian city of Iquitos, Peru. *Annals of Tropical Medicine and Parasitology* 100 Suppl 1.
8. Rocha C et al. (2009) Comparison of two active surveillance programs for the detection of clinical dengue cases in Iquitos, Peru. *American Journal of Tropical Medicine and Hygiene* 80:656-60.
9. Vazquez-Prokopec GM et al. (2009) Usefulness of commercially available GPS data-loggers for tracking human movement and exposure to dengue virus. *International Journal of Health Geographics* 8:68.
10. Paz-Soldan VA et al. (2010) Assessing and Maximizing the Acceptability of Global Positioning System Device Use for Studying the Role of Human Movement in Dengue Virus Transmission in Iquitos, Peru. *American Journal of Tropical Medicine and Hygiene* 82:723-730.
11. Morrison AC et al. (2010) Epidemiology of Dengue Virus in Iquitos, Peru 1999 to 2005: Interepidemic and Epidemic Patterns of Transmission. *PLoS Neglected Tropical Diseases* 4:e670.
12. Forshey BM et al. (2010) Arboviral etiologies of acute febrile illnesses in Western South America, 2000-2007. *PLoS Neglected Tropical Diseases* 4:e787.
13. Forshey BM et al. (2009) Dengue virus serotype 4, northeastern Peru, 2008. *Emerging Infection Diseases* 15:1815-8.
14. Siler JF, Hall MW, and Hitchens AP (1925) Results obtained in the transmission of dengue fever. *JAMA: The Journal of the American Medical Association* 84:1163-1172.
15. Kuno G (1995) Review of the Factors Modulating Dengue Transmission. *Epidemiologic Reviews* 17.
16. Chadee DD, and Martinez R (2000) Landing periodicity of Aedes aegypti with implications for dengue transmission in Trinidad, West Indies. *Journals of Vector Ecology* 25:158-163.
17. Yasuno M, and Tonn RJ (1970) A study of biting habits of Aedes aegypti in Bangkok, Thailand. *Bulletin Of The World Health Organization* 43:319-25.
18. Stoddard ST et al. (2009) The role of human movement in the transmission of vector-borne pathogens. *PLoS Neglected Tropical Diseases* 3:e481.
19. Krebs CJ (1989) *Ecological Methodology* (Harper & Row, New York).
20. Smith GM, Saveliev AA, Walker N, Ieno EN, and Zuur AF (2009) *Mixed effects models and extensions in ecology with R* (Springer).
21. Burnham KP, and Anderson DR (2002) *Model selection and multimodel inference: a practical information-theoretic approach* (Springer Verlag).
22. Baddeley A, and Turner R (2005) Spatstat: an R package for analyzing spatial point patterns. *Journal of Statistical Software* 12:1-42. Available at: www.jstatsoft.org.
23. Bates D, and Maechler M (2010) lme4: Linear mixed-effects models using S4 classes. *lme4: Linear mixed-effects models using S4 classes*. Available at: http://CRAN.R-project.org/package=lme4.
24. R Development Core Team (2010) R: A Language and Environment for Statistical Computing. *R: A Language and Environment for Statistical Computing.* Available at: http://www.R-project.org.
25. Wood SN (2006) *Generalized Additive Models: An Introduction with R* (Chapman and Hall/CRC).

26. Venables WN, and Ripley BD (2002) *Modern applied statistics with S* (Springer, New York)Fourth. Available at: http://www.stats.ox.ac.uk/pub/MASS4.

Supplementary Tables

Table S1: Summary of results for all participants (index cases included)

Season	Type	Clusters	Participants[Δ]	Febrile[a]	Acute[b]	Recent[c]	Pre-exposed[d]	Negative
2008-9	**DENV+**	12	214	21	48	28	2	136
	DENV-	5	77	0	4	6	1	66
	ND[Ω]	1	25	1	1	5	0	19
		18	***316***	***22***	***53***	***39***	***3***	***221***
2009-10	**DENV+**	11	182	16	29	12	15	126
	DENV-	20	374	3	16	31	51	276
	ND[Ω]	5	88	0	6	8	13	61
		36	***644***	***19***	***51***	***51***	***79***	***463***
	Total	***54***	***960***	***41***	***104***	***90***	***82***	***684***

Ω: Index case could not be classified as DENV4 positive or negative based on serology.
Δ: Total participation rate of 56.1%; 33 individuals participated in 2 cluster investigations.
a: Dengue positive participants presenting with fever and other symptoms.
b: Positive by PCR/IFA or IgM seroconversion between acute and convalescent sample.
c: Positive by IgM in the acute sample.
d: Positive for DENV4 antibody by $PRNT_{70}$ or confirmed positive for DENV4 by PCR/IFA in samples prior to cluster initiation; Prior antibody status was available for 45.8% of all participants.

Table S2: Summary of results of households (cluster level)

Season	Type	Clusters	Households[a]	% Acute[b]	% Recent[b]	% Pre[c]
2008-9	**DENV+**	12	40	55.00	50.00	7.41
	DENV-	5	15	26.67	33.33	11.11
	ND[Ω]	1	5	20.00	40.00	0.00
		18	***60***	***45.00***	***45.00***	***7.89***
2009-10	**DENV+**	11	38	36.84	26.32	38.10
	DENV-	21	73	17.81	27.40	57.45
	ND[Ω]	4	20	30.00	30.00	57.14
		36	**131**	**25.19**	**27.48**	**52.00**
	Total	***54***	***191***	***31.41***	***32.98***	***37.17***

Ω: Index case could not be classified as DENV4 positive or negative based on serology.
a: A total of 28 houses refused to participate in the study. 87.2% of all contact houses participated.
b: Percentage of houses including at least one individual (excluding the cluster index) with evidence of an acute, recent, or prior infection.
c: Percentage of houses with at least 1 individual with pre-exposure to dengue from houses for which evidence of prior exposure was available.

Table S3: Estimates of symptomatic:inapparent ratios for contacts within cluster studies.

Cluster type	1-θ	s.e.m	S:I
DENV+	0.179	0.047	0.218
DENV-	0.057	0.032	0.061

θ, proportion of cases that were inapparent; S:I, symptomatic to inapparent ratio.

Table S4: Summary of longitudinal cohort blood draws

Blood draw	Date		Samples	DENV-4*
1	Nov 6, 2007	Jun 9, 2008	2444	0.34%
2	Jul 31, 2008	May 6, 2009	2094	14.00%
3	Aug 19, 2009	Jan 11, 2010	2080	36.00%
4	Feb 5, 2010	Aug 13, 2010	2000	59.00%

*, Prevalence of DENV-4 neutralizing antibody based on $PRNT_{70}$ results.

Table S5: Parameter descriptions

Scale	Model		Effect	Description
Cluster	**Risk**	**y1**	DENV+	Dengue infection status of index case
	Attack rates	**y1**	acute attack rate	Proportion of previously susceptible individuals with acute infection
		y2	recent attack rate	Proportion of previously susceptible individuals with acute or recent infection
	Infestation rates	**y1**	acute infestation rate	Proportion of houses in cluster with >1 person with acute DENV infection
	Seroprevalence	**y1**	seroprevalence	Proportion of participants in a cluster with evidence of prior exposure to DENV-4
House	**Attack rates**	**y1**	acute household attack rate	Proportion of previously susceptible individuals in houses with acute DENV infections
		y2	recent household attack rate	Proportion of previously susceptible individuals in houses with acute or recent DENV infections
	Infestation risk	**y1**	Infestation risk	Presence of ≥ 1 acutely infected individual in a house
Predictors		***sigma***	sigma	Index of contact
		***sigma*$^{\Delta}$**	sigma	Index of contact calculated without index case
		home_ar	home attack rate	Attack rate in the home of the index case
		exposure	exposure	Exposure outside the house as the product of number of visits and the number of acute infections.
		site_inf	Infestation rate	Proportion of sites in activity space with >1 acute DENV infection
		cluster_a	acute infections	Total number of acutely infected contacts in activity space
		mnc	contacts	Mean number of contacts per house
		cscale	cluster scale	Summed pairwise distances between contact locations
		nsites	number of sites	Total number of sites included in cluster investigation
		age	age + age^2	
		sex	sex	
		pre	pre-exposure	Number of contacts with previous exposure to DENV-4.
		nbhd	neighborhood	TA or MY
		season	transmission season	2008-9 or 2009-10
		ctype	cluster type	DENV+ or DENV-. DENV- was the reference class.
		cluster	cluster number	ID of individual cluster. Random effect in house level models
		par	participants	Participants per house or cluster
		distance	log10(distance)	Euclidean distance in meters from index home to contact sites.

Table S6: AIC table of models for individual risk

Model	Parameters	K	AICc	Δ_i	w_i	log(*L*)
13	***sigma***$^{\circ\circ}$***+pre***$^{\#}$***+age***	5	43.34	0.00	0.42	-15.95
12	***sigma+pre***	3	43.69	0.35	0.35	-18.57
14	***sigma***	2	45.38	2.05	0.15	-20.56
9	***sigma+cscale+season+pre***	5	46.59	3.26	0.08	-17.58
1	***season+nbhd+season:nbhd +cscale+mnc+sigma+pre+age +sex***	11	53.49	10.16	0.00	-12.08
15	***sigma***$^{\Delta *}$***+pre***$^{+}$***+age***	5	58.01	14.67	0.00	-23.29
16	***site_inf+pre***	3	58.18	14.84	0.00	-25.82
5	***cluster_a + pre***	3	59.21	15.88	0.00	-26.33
6	***site_inf+pre+season+cscale***	5	62.17	18.83	0.00	-25.37
7	***cluster_a+site_inf***	3	63.29	19.96	0.00	-28.37
3	***home_ar***	2	65.42	22.09	0.00	-30.57
17	***home_ar+exposure***	3	66.20	22.86	0	-29.82
10	***season+nbhd+season:nbhd +cscale***	5	68.88	25.54	0.00	-28.73
11	***mnc+cscale+pre+age+sex***	7	69.04	25.71	0.00	-26.12
4	***cscale+nsites***	3	69.81	26.47	0.00	-31.63
8	***sex*age***	5	72.26	28.92	0.00	-30.41
2	***season+nbhd+season:nbhd +home_ar+age+sex***	9	72.59	29.25	0.00	-24.86

K = number of parameters, AIC_c = AIC corrected for small sample size, Δ_i = difference in AIC_c between model *i* and best model, w_i = Akaike weight, log(L) = log likelihood. Within formulae, colon indicates interaction (e.g., x1:x2), asterisk indicates both main and interaction effects (e.g., x1+x2+x1:x2). Parameter definitions in Table S5. Significance: $^{\circ\circ\circ}$ $p < 0.001$, $^{\circ\circ}$ $p < 0.01$, $^{\circ}$ $p < 0.05$, $\#$ $p < 0.1$. The best model is at the top.

Table S7: Models for acute attack rates within activity spaces (cluster scale)

Model	Parameters	K	QAICc	Δ_i	w_i	log(*L*)
4	***ctype***$^{\circ\circ\circ}$***+season***$^{\#}$	4	109.37	0.00	0.35	-50.65
8	***ctype***	3	109.68	0.31	0.30	-51.82
7	***ctype+par***	4	111.39	2.03	0.13	-51.67
5	***ctype+nbhd***	4	111.68	2.32	0.11	-51.81
2	***season*nbhd*ctype***	9	112.96	3.59	0.06	-47.34
6	***ctype+cscale+par***	5	113.43	4.06	0.05	-51.67
1	***season*nbhd*ctype+cscale+par***	11	116.78	7.41	0.01	-47.19
3	***season*nbhd***	5	118.60	9.23	0.00	-54.25

Significance: $^{\circ\circ\circ}$ $p < 0.001$, $^{\circ\circ}$ $p < 0.01$, $^{\circ}$ $p < 0.05$, $\#$ $p < 0.1$; $\hat{c}$ estimate = 1.40

Table S8: Models for recent attack rates within activity spaces (cluster scale)

Model	Parameters	K	QAICc	Δ_i	w_i	log(*L*)
4	***ctype***$^{\circ\circ}$***+season***$^{\circ}$	4	108.16	0.00	0.33	-50.05
8	***ctype***	3	108.32	0.16	0.30	-51.14
7	***ctype+par***	4	109.72	1.56	0.15	-50.83
5	***ctype+nbhd***	4	110.33	2.17	0.11	-51.14
6	***ctype+cscale+par***	5	111.16	3.00	0.07	-50.54
3	***season*nbhd***	5	113.39	5.24	0.02	-51.66
2	***season*nbhd*ctype***	9	115.80	7.64	0.01	-48.78
1	***season*nbhd*ctype+cscale+par***	11	118.70	10.54	0.00	-48.17

Significance: $^{\circ\circ\circ}$ $p < 0.001$, $^{\circ\circ}$ $p < 0.01$, $^{\circ}$ $p < 0.05$, $\#$ $p < 0.1$; $\hat{c}$ estimate = 2.03

Table S9: Models of house infestation rates within activity spaces (cluster scale)

Model	Parameters	K	QAICc	Δ_i	w_i	log(*L*)
7	***ctype***$^{\circ\circ\circ}$	3	108.69	0.00	0.35	-51.28
4	***ctype+season***	4	108.91	0.22	0.31	-50.35
6	***ctype+cscale***	4	110.29	1.61	0.15	-51.04
5	***ctype+nbhd***	4	110.62	1.93	0.13	-51.20
2	***season*nbhd*ctype***	9	113.27	4.59	0.03	-47.14
1	***season*nbhd*ctype+cscale***	10	114.43	5.75	0.02	-46.61
3	***season*nbhd***	5	116.69	8.00	0.01	-53.18

Significance: $^{\circ\circ\circ}$ $p < 0.001$, $^{\circ\circ}$ $p < 0.01$, $^{\circ}$ $p < 0.05$, $\#$ $p < 0.1$; $\hat{c}$ estimate = 1.16

Table S10: Models of DENV-4 antibody prevalence within activity spaces (cluster level)

Model	Parameters	K	AICc	Δ_i	w_i	log(*L*)
2	***season***$^{\circ\circ\circ}$***+nbhd***$^{\#}$	3	116.37	0.00	0.52	-55.15
5	***season***	2	117.63	1.26	0.28	-56.80
3	***ctype+season***	3	118.57	2.20	0.17	-56.25
1	***season*nbhd*ctype***	8	121.81	5.44	0.03	-52.73
4	***ctype+nbhd***	3	146.62	30.26	0.00	-70.28

Significance: $^{\circ\circ\circ}$ *p < 0.001,* $^{\circ\circ}$ *p < 0.01,* $^{\circ}$ *p < 0.05,* # *p < 0.1;* $\hat{c}$ estimate = 1

Table S11: Models of household attack rates (house scale)

Model	Parameters	K	QAICc	Δ_i	w_i	log(*L*)
3	***ctype***$^{\circ\circ\circ}$***+season***	4	231.47	0.00	0.33	-111.70
6	***ctype***	3	231.52	0.05	0.32	-112.74
5	***ctype+htype***	4	232.60	1.13	0.19	-112.27
4	***ctype+nbhd***	4	233.50	2.03	0.12	-112.72
2	***season+nbhd+ctype*htype***	7	235.97	4.50	0.03	-110.90
1	***season*nbhd*ctype*htype***	23	250.57	19.10	0.00	-101.41

All models include offset term for whether the house participated in the longitudinal cohort. Significance: $^{\circ\circ\circ}$ *p < 0.001,* $^{\circ\circ}$ *p < 0.01,* $^{\circ}$ *p < 0.05,* # *p < 0.1;* $\hat{c}$ estimate = 1.17

Table S12: Models for household infestation risk (house scale)

Model	Parameters	K	AICc	Δ_i	w_i	log(*L*)
5	***ctype***$^{\circ\circ\circ}$***+distance***	3	146.88	0.00	0.37	-70.35
7	***ctype***	2	147.58	0.71	0.26	-71.75
6	***ctype+season***	3	148.05	1.17	0.21	-70.93
2	***ctype*distance***	4	148.77	1.90	0.15	-70.23
1	***season*ctype*distance+nbhd***	9	154.75	7.88	0.01	-67.65
3	***season*distance***	4	159.13	12.26	0.00	-75.41
4	***nbhd*distance***	4	164.84	17.96	0.00	-78.26

Significance: $^{\circ\circ\circ}$ *p < 0.001,* $^{\circ\circ}$ *p < 0.01,* $^{\circ}$ *p < 0.05,* # *p < 0.1*

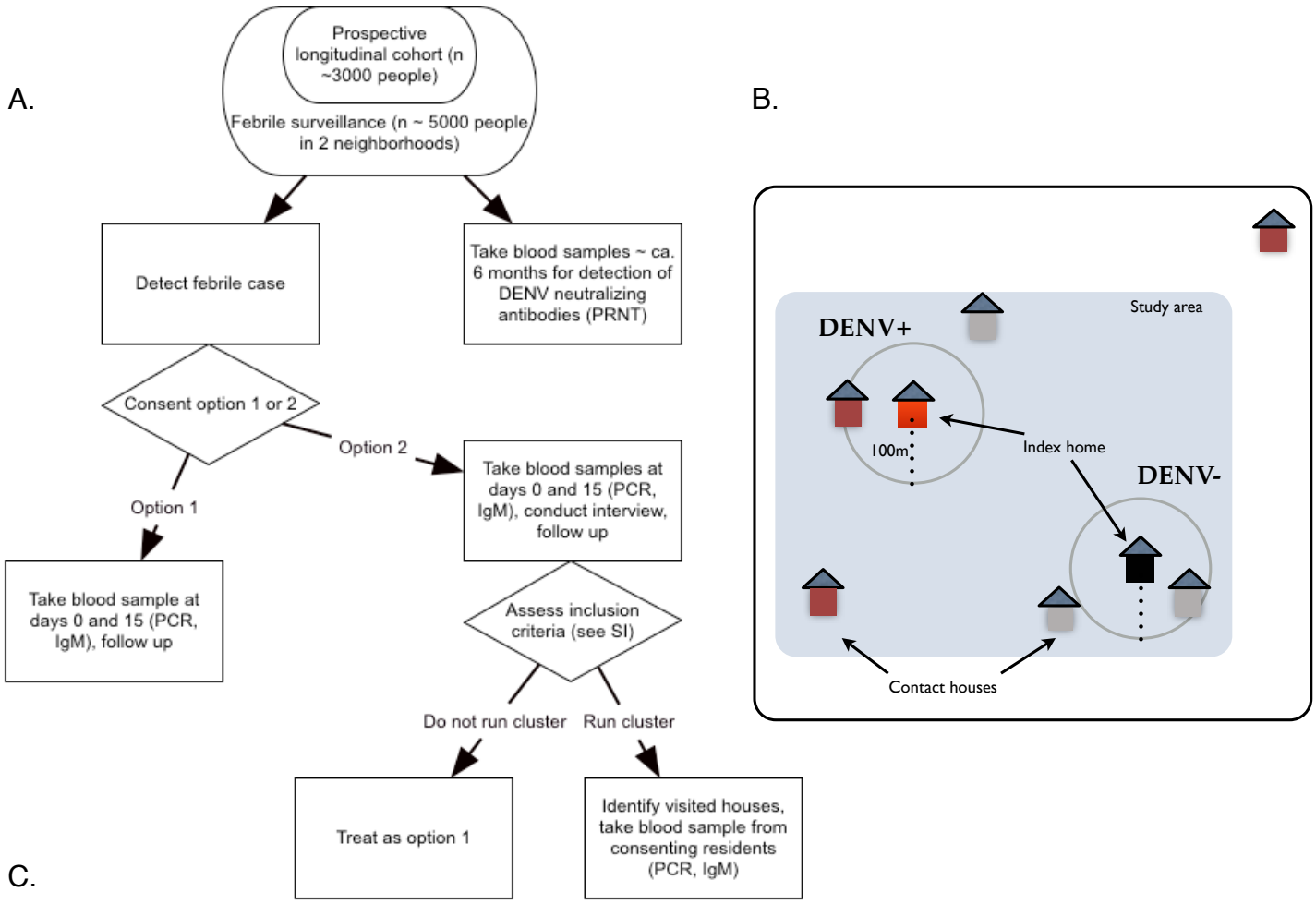
Table S13: Recent attack rate estimates for activity spaces (cluster level)

Season	Cluster type	AR	s.e.m.
2008-9	**DENV+**	29.90%	0.030
	DENV-	19.90%	0.034
2009-10	**DENV+**	22.00%	0.029
	DENV-	14.00%	0.018

Table S14: Parameter estimates from seasonally-forced transmission model.

Parameter	mle (95% CI)	posterior mean (95% CI)
$C^M = C^T$	2.09×10^{-4} $(7.02 \times 10^{-5}, 4.23 \times 10^{-4})$	1.77×10^{-4} $(0, 4.14 \times 10^{-4})$
β_1^M	1.01×10^{-3} $(1.52 \times 10^{-4}, 2.11 \times 10^{-3})$	1.12×10^{-4} $(2.58 \times 10^{-5}, 4.23 \times 10^{-4})$
$\beta_1^T = 0$	NA	NA
$\beta_2^M = 0$	NA	NA
β_2^T	1.70×10^{-3} $(8.06 \times 10^{-4}, 2.49 \times 10^{-3})$	1.64×10^{-3} $(5.44 \times 10^{-4}, 2.66 \times 10^{-3})$
$\beta_3^M = \beta_3^T = 0$	NA	NA
$\beta_4^M = \beta_4^T$	1.37×10^{-3} $(7.75 \times 10^{-4}, 2.21 \times 10^{-3})$	1.66×10^{-3} $(7.01 \times 10^{-4}, 2.80 \times 10^{-3})$
$p_1^{(3)M} = p_1^{(3)T} = p_2^{(3)T}$	0.158 $(0.105, 0.220)$	0.160 $(0.092, 0.231)$
$p_2^{(3)M} = 0$	NA	NA
$\alpha^M = \alpha^T$	131.947 $(96.656, 172.506)$	123.221 $(76.450, 185.651)$

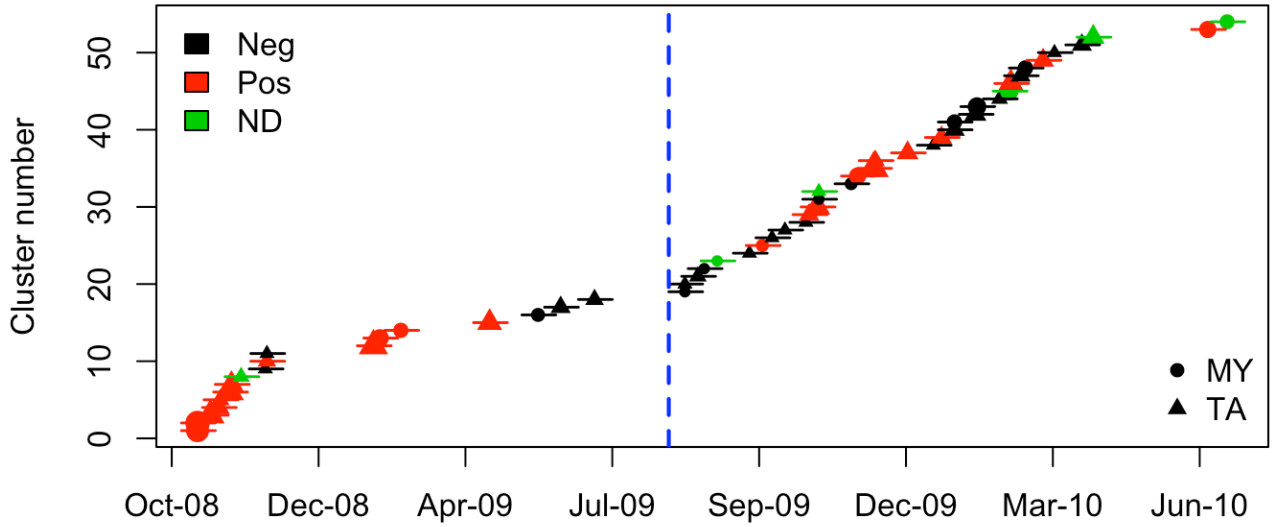
Supplementary Figures



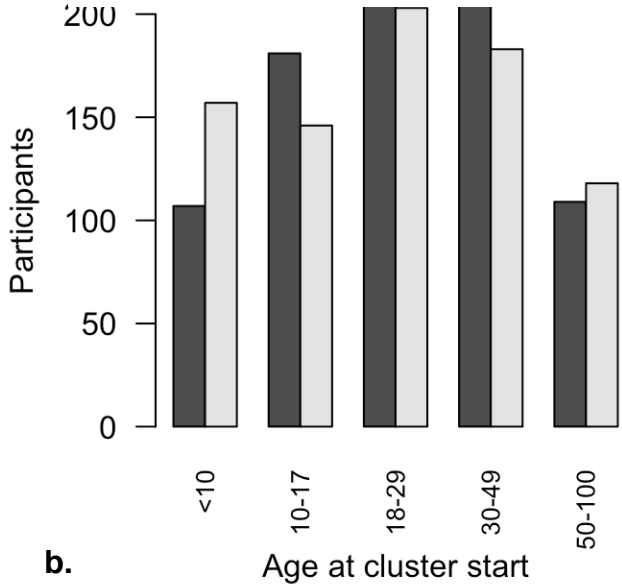
Cluster invest.	Cluster type†	House	Participant	Result
CI_1	DENV+	House_1*	Par_1.1.01	Acute D4
CI_1	DENV+	House_1	Par_1.1.02	Neg
CI_1	DENV+	House_2	Par_1.2.01	Recent DENV
CI_1	DENV+	House_2	Par_1.2.02	Acute D4
CI_1	DENV+	House_3	Par_1.3.01	Acute D4
…				
CI_2	DENV-	House_1	Par_2.1.01	Neg
CI_2	DENV-	House_2	Par_2.2.01	Immune
CI_2	DENV-	House_2	Par_2.2.02	Neg
CI_2	DENV-	House_2	Par_2.2.03	Neg
CI_2	DENV-	House_2	Par_2.2.04	Recent DENV

*Index house.

Figure S1: Overview of study design. A) Flow chart of activities. Febrile cases were detected in two study areas, triggering case and control cluster investigations. B) Geographic representation of case-control cluster studies. Study houses could occur throughout urban Iquitos. C) Example of data collected for analysis showing two cluster investigations. Infection rates were estimated at the level of clusters and households. See Text S1 for additional details.



a.



b.

Figure S2. Timing and demographics. *a:* Timing of all cluster investigations run with type and neighborhood of origin indicated. Indeterminate clusters (ND) are included in green. The size of symbols is commensurate with attack rates in the clusters. Blue line demarcates season 1 from season 2. *b:* Age and sex distribution of all participants (index cases and contacts). Black, females; gray, males.

Figure S3. Demographic composition of index cases. *a*: Age distribution. *b*: Self-reported occupation.

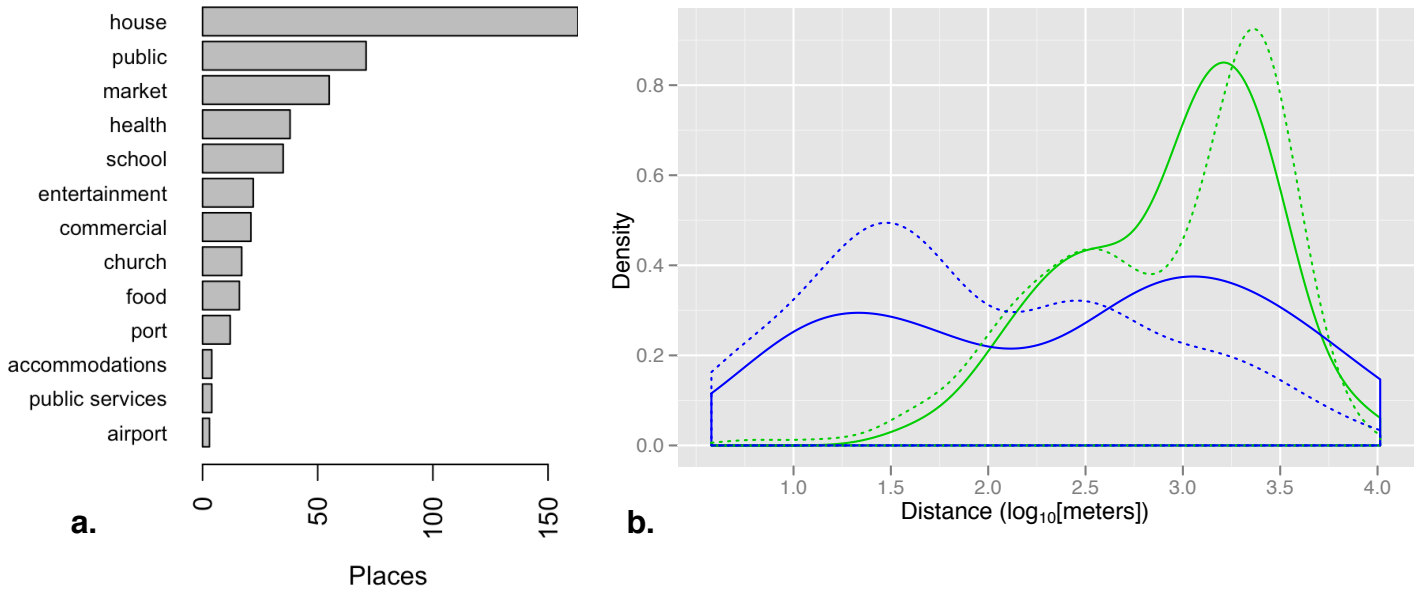
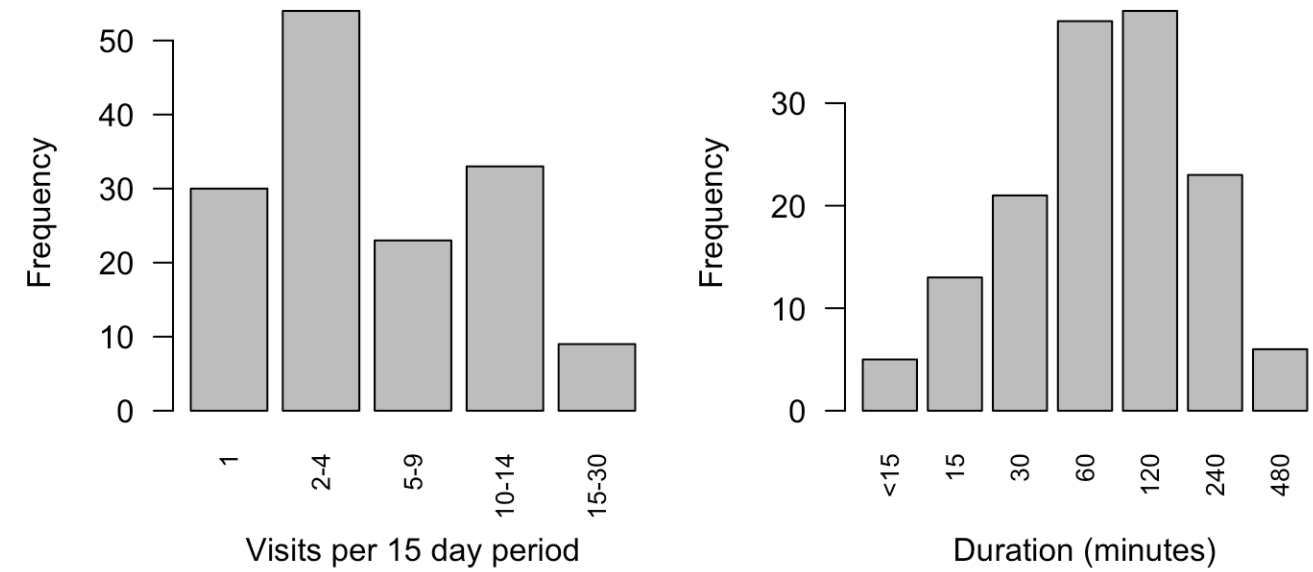


Figure S4. Movement patterns of index cases. *a*: Types of places visited organized into broad categories. Note a few places appear in 2 categories because they had distinct uses (e.g., one house also served as a church). Also, schools are on holiday between mid-December and early March each year, which corresponds with the second half of the transmission season. Public places: plazas, parks, and other recreation areas. Markets: largely open-air markets. Health: hospitals and clinics. Entertainment, bars. Food, restaurants and eateries. *b*: Distance kernels (from the home) by neighborhood of residence (MY, solid line; TA, dashed line) and type of place visited (houses, blue; all other places, green). These data are from all 54 cluster investigations, of which twice as many index cases were from TA (36) compared to MY (18).



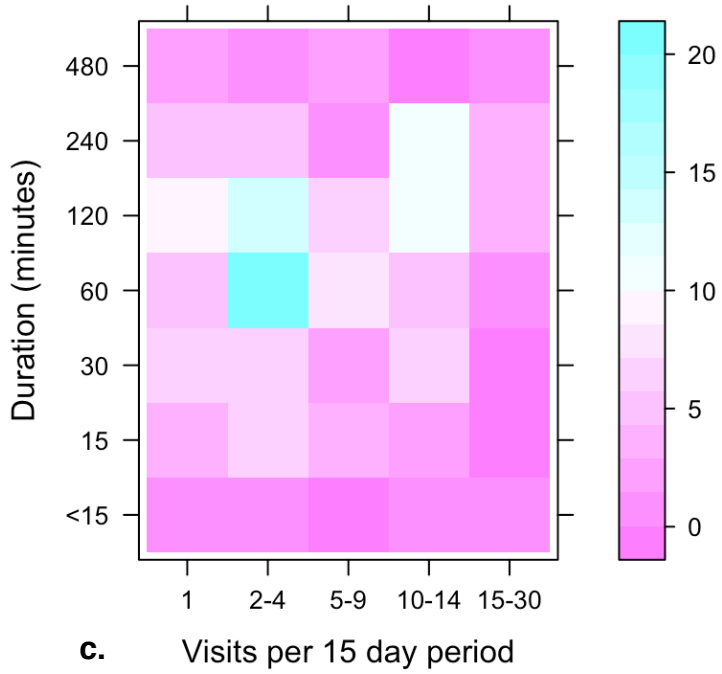


Figure S5. Temporal aspects of visits to other houses reported by index cases. *a*: Number of visits to a house in the previous 15 days (median = 4 visits). *b*: Distribution of the duration (in minutes) of each visit (median = 90 minutes). *c*: Level plot showing relationship between number and duration of visits. The most common response was 2-4 visits of an hour over the previous 15 days.

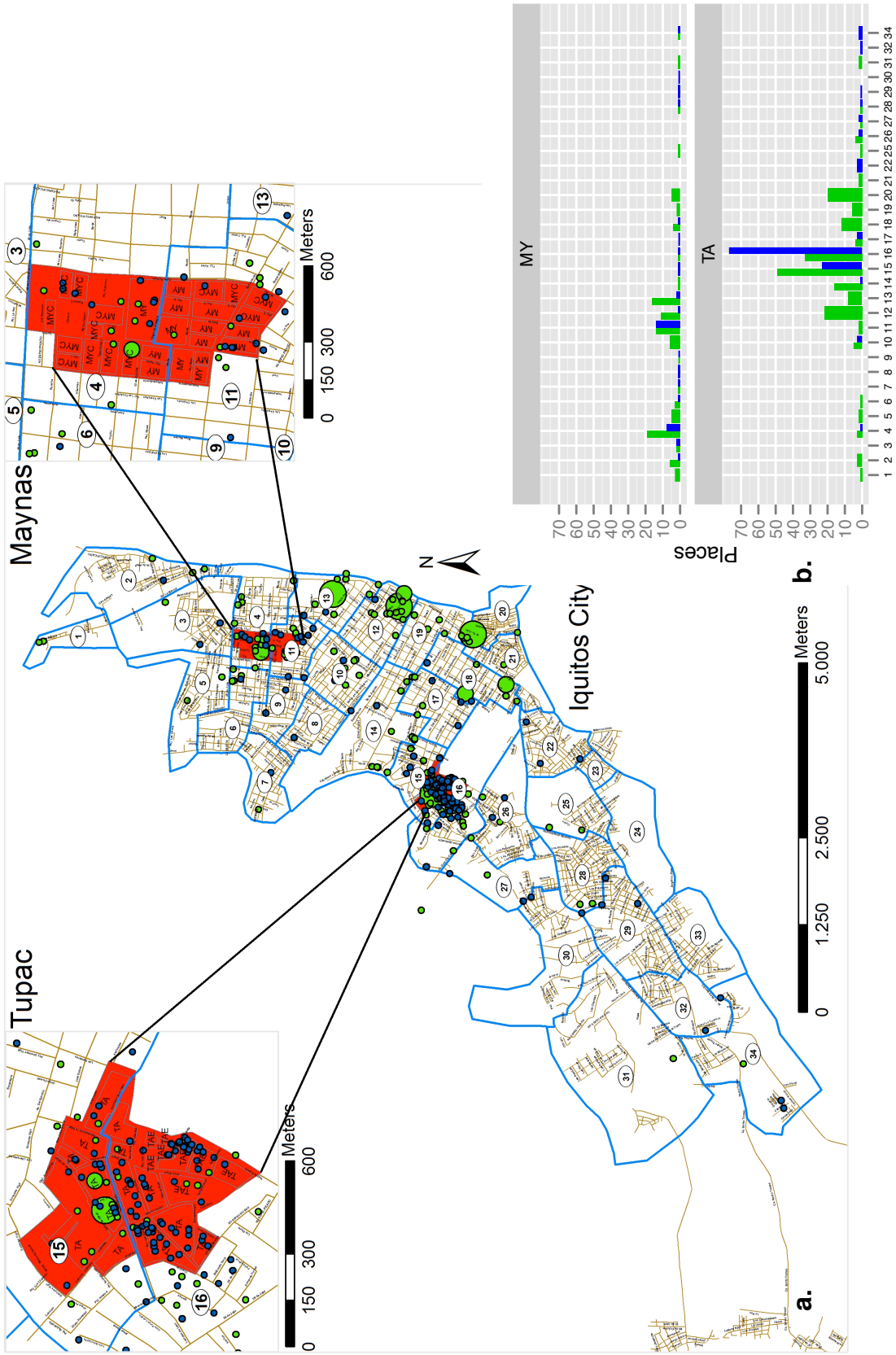


Figure S6. Movement patterns of index cases, including all places reported as visited in the previous 15 day period (excluding the home). *a*: Map of Iquitos showing the 2 study neighborhoods (Tupac = TA, Maynas = MY) and the health sectors (outlined in blue). *b*: The number and type of place visited (houses, blue; all other places, green) in different health sectors within the city varied by neighborhood of origin, with limited overlap. Health sectors are roughly arranged from North to South and encompass approximately 2500 houses each. The downtown area falls in sector 12 and the principal city market is in sector 20. These data are from all 54 cluster investigations, of which twice as many index cases were from TA (36) compared to MY (18). Residents of TA on average visited 1 more house than residents of MY (3.39 vs 2.28).

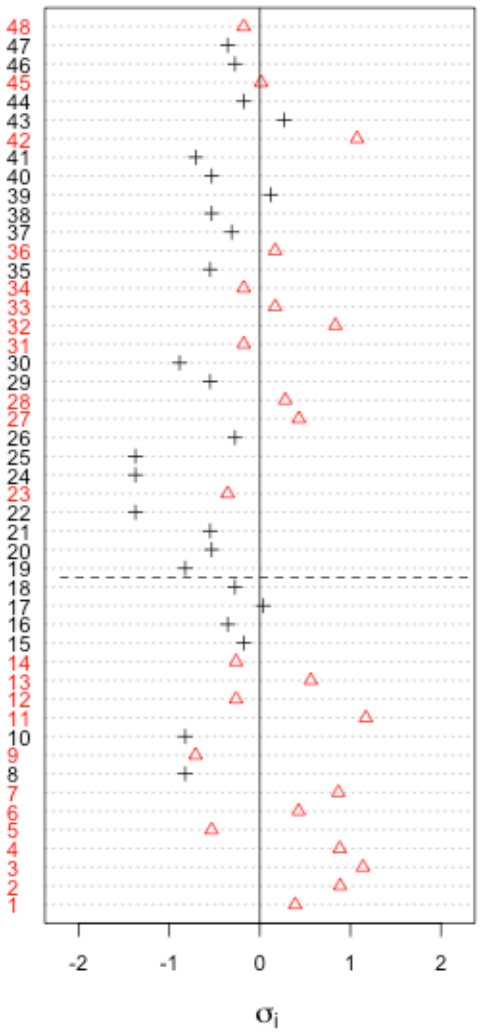


Figure S7. Contact rates. σ_i for the 48 positive or negative cluster investigations in chronological order from bottom to top. Dashed line demarcates first and second seasons. Red triangles, DENV+ clusters, black crosses, DENV- clusters.

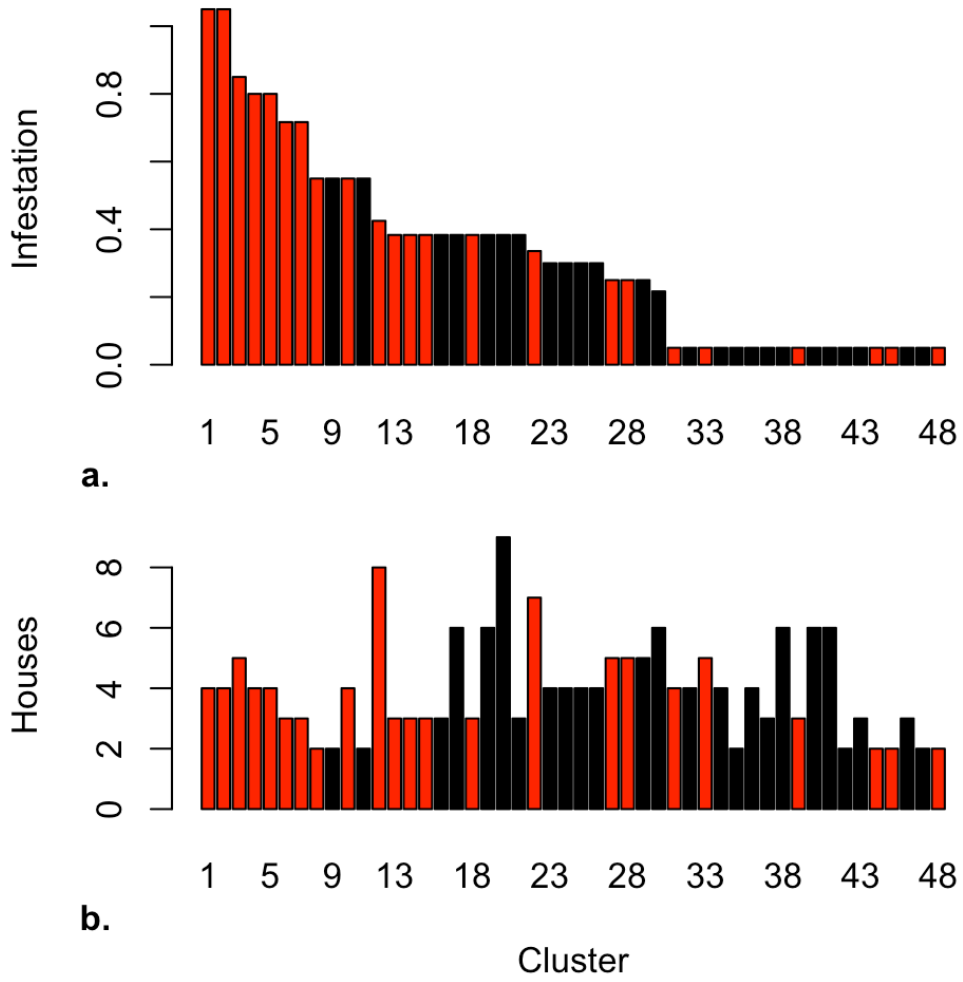


Figure S8. Transmission patterns across all clusters. *a*: Infestation rates (proportion of houses in cluster with ≥1 acute DENV infection). *b*: Number of houses in each cluster investigation. Values are organized in order of decreasing infestation rate and colored by type of cluster (red, DENV+; black, DENV-). Note 0.05 was added to all infestation rates to visualize clusters with no acute infections in panel *a*.

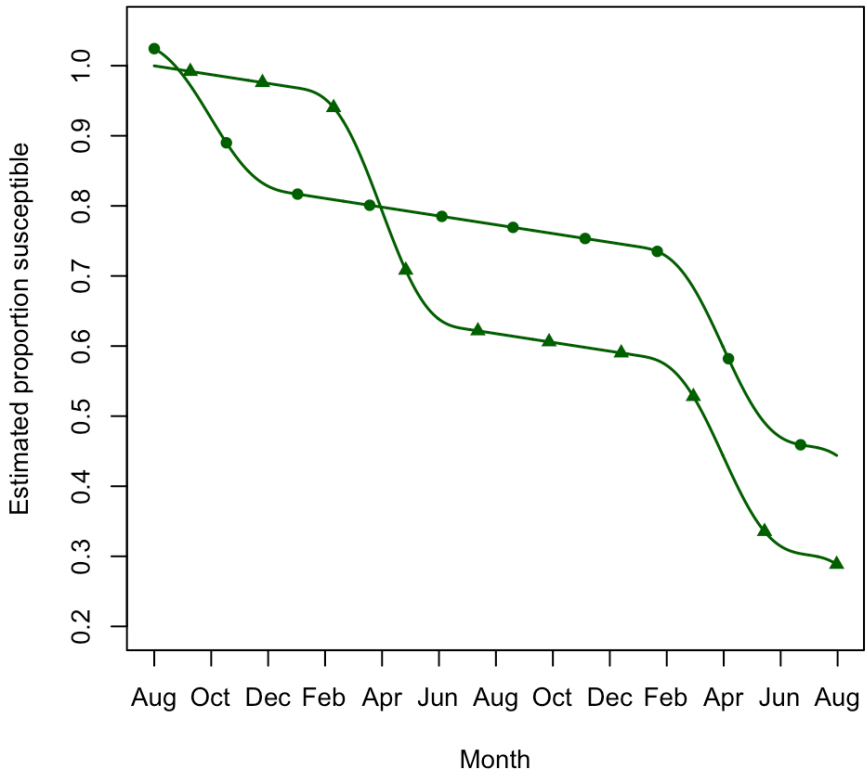


Figure S9. **Susceptibility to DENV-4.** Population susceptibility estimated with the transmission model (Eq. S7). Circle, MY neighborhood, triangle, TA neighborhood.

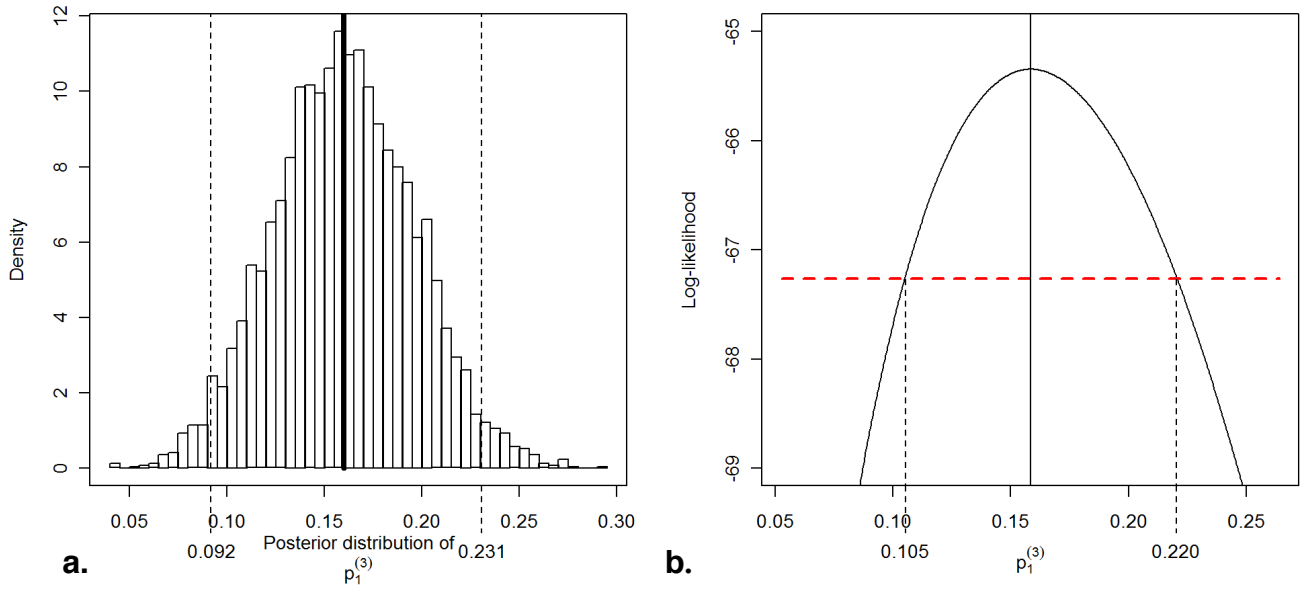


Figure S10. **Estimates of $p^{(3)}$, the within-cluster infection probability**. *a,* Posterior distribution. Dashed lines indicate 95% CI. *b,* Profile likelihood based 95% confidence interval, shown by dashed vertical lines.

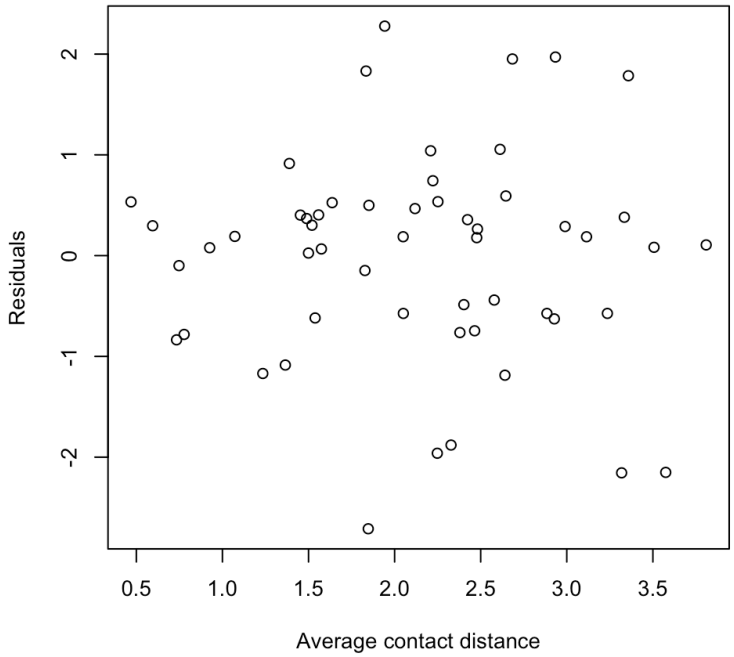


Figure S11. Effect of distance from index home on transmission. Residuals from transmission model (observed - predicted) plotted against the average distances ($\log_{10}$[meters]) of cluster contacts (based on their home location) from the index case.