MOLECULAR ECOLOGY

Human facilitated metapopulation dynamics in an emerging pest species, *Cimex lectularius*

Journal:	*Molecular Ecology*
Manuscript ID:	MEC-13-1080
Manuscript Type:	Original Article
Date Submitted by the Author:	03-Oct-2013
Complete List of Authors:	Fountain, Toby; University of Sheffield, Animal and Plant Sciences; University of Helsinki, Department of Biosciences Horsburgh, Gavin; University of Sheffield, Animal and Plant Sciences Duvaux, Ludovic; University of Sheffield, Animal and Plant Sciences Reinhardt, Klaus; University of Sheffield, Animal and Plant Sciences; University of Tübingen, Institute for Evolution and Ecology Butlin, Roger; University of Sheffield, Animal and Plant Sciences
Keywords:	Ecological Genetics, Population Genetics - Empirical, Population Dynamics, Inbreeding, Insects

HUMAN FACILITATED METAPOPULATION DYNAMICS IN AN EMERGING PEST SPECIES, *CIMEX LECTULARIUS*

[1,2]Toby Fountain, [1]Gavin Horsburgh, [1]Ludovic Duvaux, [1,3]Klaus Reinhardt, [1]Roger K. Butlin

[1]Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield, S10 2TN. UK

[2]Department of Biosciences, PO Box 65 (Viikinkaari 1), FI-00014 University of Helsinki, Finland

[3]Institute for Evolution and Ecology, University of Tübingen, Auf der Morgenstelle 28, D-72076 Tübingen. Germany

Keywords: ABC analysis, *Cimex lectularius*, genetic structure, metapopulation dynamics, microsatellites, pest management

Corresponding author: Toby Fountain, Department of Biosciences, PO Box 65 (Viikinkaari 1), FI-00014 University of Helsinki, Finland Email: toby.fountain@helsinki.fi

Running title: Metapopulation dynamics of *Cimex lectularius*

ABSTRACT

The number and demographic history of colonists can have dramatic consequences for the way genetic diversity is distributed and maintained in a metapopulation. Here, using 21 newly characterised microsatellite markers and Approximate Bayesian Computation (ABC) in a metapopulation framework we estimate the number and genetic composition of founders in a re-emerging pest species, the common bed bug (*Cimex lectularius*). The bed bug's close association with humans has led to metapopulation dynamics of frequent local extinction and colonisation. Pest control limits the lifespan of sub-populations, causing frequent local extinctions, and human-facilitated dispersal allows the colonisation of empty patches. Founder events often result in drastic reductions in diversity and an increased influence of genetic drift. Coupled with restricted migration this can lead to rapid population differentiation. We therefore predicted strong population structuring. We found very limited diversity within-infestations but high degrees of structuring across the city of London, with extreme levels of genetic differentiation between infestations (F_{ST} = 0.59). ABC results suggest a common origin of founders and that the numbers of colonists were low; implying that even a single mated female is enough to found a new infestation successfully. These patterns of colonisation are close to the predictions of the propagule pool model, where all founders originate from the same parental infestation. These results not only formalise the applicability of a metapopulation approach with this system, but also provide a valuable resource for the future targeted control of bed bug infestations.

INTRODUCTION

With increased travel and connectivity there are now many opportunities for the human-facilitated dispersal of organisms, with disease vectors and pests of economic importance presenting a particular concern (Estoup *et al.* 2004; Grapputo *et al.* 2005; Tatem *et al.* 2006; Niggemann *et al.* 2009; Lawson Handley *et al.* 2011). Passive dispersal, coupled with high population turnover, can lead to organisms existing as highly structured metapopulations (De Meester *et al.* 2002; Haag *et al.* 2005; Walser & Haag 2012) with discrete breeding groups, frequent and independent local extinctions, and potential for patches to be recolonised (see Hanski 1999).

The population dynamics of metapopulations greatly affect how genetic diversity is distributed, both within and between local populations (Slatkin 1977; Wade & McCauley 1988; Hastings and Harrison 1994). Founder events can shift populations out of equilibrium by reducing genetic diversity and increasing genetic drift. Coupled with restricted migration this can lead to rapid genetic differentiation of sub-populations (Wade & McCauley 1988; Whitlock & McCauley 1990). Frequent local extinction events can make differentiation even more extreme by limiting the ability of gene flow to equalise allele frequencies (Whitlock & McCauley 1990; Hastings & Harrison 1994; Pannell & Charlesworth 1999; Ray 2001).

Theoretical studies predict that the number and origin of colonists of each sub-population has a strong influence on the degree of sub-population differentiation. When colonists of any given sub-population have originated from a unique parental population (propagule

pool model), differentiation between sub-populations is always predicted to increase relative to the equilibrium with no extinction. This is in tandem with a reduction in neutral genetic variation both within sub-populations and throughout the entire metapopulation (Wade & McCauley 1988; Pannell & Charlesworth 1999). In contrast, a mixed origin of colonists (migrant pool model) may lead to an increase or decrease in differentiation (Wade & McCauley 1988). The numbers of founders and their relatedness to each other therefore directly influence the genetic variance among sub-populations, with kin-structured colonisation and subsequent inbreeding leading to substantial differentiation (Whitlock 1992; Ingvarsson & Olsson 1997; Ingvarsson 1998; Johannesen & Lubin 1999; Torimaru *et al.* 2007).

Metapopulation dynamics are not only important for understanding many evolutionary processes but also have practical implications for conservation (Frankham et al. 2002; Ray 2001), control of invasive species (Lawson Handley *et al.* 2011) and integrated pest management (Collins *et al.* 2000; Rinkevich *et al.* 2007; Yakob *et al.* 2008). For example, selection for resistance alleles could reduce the chance of a sub-population going extinct after a control treatment, thus increasing the period during which it acts as a source of migrants. This would facilitate the spread of resistance alleles into new sub-populations and severely hamper control. Knowledge of the nature of population structure would help predict the spread of resistance alleles (Churcher *et al.* 2008).

The common bed bug (*Cimex lectularius*) is re-emerging as a significant economic and public health pest, precipitated by a sudden global resurgence in its populations (Boase 2001;

Doggett *et al.* 2004; Kilpinen *et al.* 2008; Potter *et al.* 2008; Richards *et al.* 2009). The causes of this sudden population expansion have remained a mystery whose resolution has been hampered by a lack of research on the bed bug's basic population and dispersal biology (Reinhardt & Siva-Jothy 2007). Bed bugs are flightless, obligate blood-sucking insects that can form infestations comprising thousands of individuals (Wang *et al.* 2010; Reinhardt *et al.* 2010). Infestations typically consist of aggregations of individuals located in discrete refugia. Bed bugs walk to the host and return to a refuge when feeding is complete (Reinhardt & Siva-Jothy 2007). Being flightless, individuals can only move actively over limited distances and much of their recent spread is attributed to long-distance passive dispersal facilitated by human movement (Doggett *et al.* 2004; Reinhardt & Siva-Jothy 2007; Potter *et al.* 2008; Szalanski *et al.* 2008).

In a bed bug metapopulation, human dwellings form habitat patches and it is expected that limited numbers of individuals found new infestations. Bed bugs have been observed actively dispersing throughout buildings (Doggett & Russell 2008; Wang *et al.* 2010) and it is likely that this not only accounts for propagation of infestations within buildings but is also the mechanism by which individuals move into portable items, leading to passive dispersal. Research on the composition of founders has been limited and recent studies have given contrasting results. For example, Szalanski et al. (2008), found up to six mitochondrial haplotypes within a single infestation and Booth et al. (2012) showed evidence of multiple introductions within an apartment complex. These results are consistent with a migrant pool model of colonisation, with a more genetically diverse group of founders. In contrast, other infestations have been shown to have very limited within-infestation diversity (Booth et al.

2012, Saenz et al. 2012), which more closely follows the predictions of a propagule pool model. Local extinctions are especially frequent as infested properties are often treated with insecticides, giving the majority of occupied patches a relatively short lifespan. Reports of widespread insecticide resistance (Doggett *et al.* 2004; Romero *et al.* 2007; Potter *et al.* 2008) suggest that selection is further shaping the genetic diversity of bed bug metapopulations. In combination, these factors predict low diversity within and very high levels of structuring between sub-populations of bed bugs.

Despite its importance in the maintenance of genetic diversity, there has been limited research on the number and genetic composition of colonists in natural metapopulations (Gaggiotti *et al.* 2004 but see Whitlock 1992; Austin *et al.* 2010). Here we use a metapopulation framework to provide detailed insight into bed bug population dynamics across two hierarchical levels of structure. We first characterized within-infestation and city-wide genetic diversity before estimating the number and demographic history of bed bug colonists using Approximate Bayesian Computation (ABC). Therefore this study, for the first time, uses a model-based approach to test, in a robust statistical framework, two alternative hypotheses on the genetic composition of founders in a natural bed bug metapopulation. We discuss the implications of these results for the future integrated control of bed bugs.

MATERIALS AND METHODS

Sample collection

To investigate within-infestation diversity, individuals were sampled from multiple refugia in five properties. These properties are subsequently referred to as AUS (sourced from New South Wales, Australia), BIR1 and BIR2 (separate properties located within Birmingham, UK) and LON1 & LON2 (separate properties from London, UK) (Table S1). Detailed spatial surveys of properties LON1 and LON2 are described in Naylor (2012), case studies 4 and 3 respectively). These infestations were selected from available samples on the basis of having individuals sampled separately from multiple refugia, rather than based on their geographic location.

To assess diversity at the city scale, 13 infestations from across London, UK, were sampled (see Figure S1, Table 3 for names and spatial locations). Pest control operatives, who obtained individuals prior to the treatment of affected properties, provided the majority of samples. Once received, all samples were stored in screw-topped rubber sealed microfuge tubes containing 1.5 ml of absolute ethanol (analytical reagent grade) at room temperature.

DNA extraction and genotyping

The ammonium acetate precipitation method described by Nicholls *et al.* (2000) was used to extract DNA. Individuals were then genotyped using 21 newly isolated microsatellite markers. The microsatellite loci were isolated from either a microsatellite-enriched genomic library or a recently available transcriptome assembly (O. Otti and K. Reinhardt, in prep; see

Supporting Information for a description of microsatellite isolation and characterisation). All 21 loci used were confirmed as autosomal by the observed presence of heterozygotes in male and female individuals. The same PCR conditions were used as described for primer testing (see Supporting Information), with individuals genotyped at 19 autosomal loci for the within-infestation study and at 21 autosomal loci for the city-wide study. Amplified products were analysed using an ABI3730 48-well capillary DNA analyser and allele sizes were assigned using GENEMAPPER v.3.7 (Applied Biosystems, California). Sequences were searched against the NCBI nr nucleotide database using BLASTn, which confirmed that these markers did not overlap with previously published bed bug microsatellites (Booth et al. 2012).

Genetic diversity

For each dataset descriptive summary statistics including number of alleles and expected (H_E) and observed heterozygosities (H_O) were obtained using Microsatellite Analyser version 4.05 (Dieringer & Schlötterer 2003). Allelic richness was calculated using F_{STAT} version 2.9.3.2 (Goudet 1995). We tested for deviations from Hardy-Weinberg equilibrium (HWE) and estimated the frequency of null alleles with Cervus version 3.0.3 (Kalinowski *et al.*, 2007). Evidence of linkage disequilibrium was assessed using GENEPOP version 4.1.0 (Raymond & Rousset 1995; Rousset 2008). For analyses of deviation from HWE and evidence of linkage disequilibrium, a Bonferroni correction was applied to allow for multiple tests (Rice 1989).

Population structure

We used F_{STAT} to calculate global F_{IS} and F_{ST} values (Weir & Cockerham 1984), both within and among infestations. Values were jack-knifed over loci to give means and standard errors and were bootstrapped over loci to give 95% confidence intervals. 10,000 permutations were used to generate significance values.

The within-city dataset was tested for isolation by distance amongst individuals in SPAGeDi ver 1.3 (Hardy & Vekemans 2002) using the kinship coefficient (Loiselle *et al.* 1995). Distance was partitioned into 10 intervals, with a uniform number of pairwise comparisons per interval. The mean distance value of each interval was log transformed (Rousset 1997). We used 10,000 permutations to test if the slope of the relationship between geographic and genetic distance was significantly negative.

For both datasets we performed a Discriminant Analysis of Principal Components (DAPC) (Jombart *et al.* 2010) using the ADEGENET 1.3-4 (Jombart 2008) package in R (R Core Team 2012) to examine evidence for genetic clusters, using infestation as a grouping prior. DAPC is an ideal clustering method for this dataset as it does not make some commonly required assumptions (e.g. Hardy-Weinberg equilibrium) (Jombart *et al.* 2010), which are unlikely to hold for bed bug infestations. The first step in DAPC is to transform the raw data into principal components (PCs). There is a subsequent trade-off in the number of retained PCs, with a higher number of PCs increasing the ability to discriminate between groups at the cost of the reduced stability of membership probabilities (Jombart *et al.* 2010). We used *a*-score as a measure for judging the optimal number of retained principal components. The *a*-

score is the difference between the proportions of successful observed discriminations and values obtained from random discrimination. This was calculated with 100 permutations for each increasing number of retained principal components using the *optim.a.score* function in ADEGENET. Due to the low number of sampled individuals in each group, we were conservative with the number of retained principal components (PCs) but in both cases the number of retained PCs still incorporated ≥75% of the variance in the data. The *dapc* function was then used to perform the clustering analysis and results are presented as ordination plots.

Approximate Bayesian Computation (ABC) analysis

ABC allows likelihood-free inference in complex scenarios with many parameters by comparing summary statistics from observed data to summary statistics from data simulated using various prior distributions (Beaumont 2010). In population genetics, where it is becoming a widely used tool, ABC analysis usually involves the construction of historical population models, simulation of many data sets using a coalescent approach and comparison of simulated to observed data using summary statistics describing genetic diversity. This provides a rigorous framework to compare different demographic models and then infer demographic parameters of interest (for a review of the global ABC procedure see Csilléry *et al.* (2010)).

In order to understand the dynamics of a bed bug metapopulation, we simulated the sampling of 20 microsatellites from 13 infestations under two demographic scenarios, mimicking the two theoretical models of metapopulation colonisation (Figure 1). Note that

from our 21 original microsatellite, the locus Cle001 was omitted, due to a large range of allele sizes, which reduced our ability to fit the models. Firstly, a migrant pool model was constructed where random individuals from the metapopulation found new infestations (scenario one) and, secondly, a propagule pool model was designed such that colonists all come from the same source infestation (scenario two) (Figure 1). In both cases, we would expect a severe bottleneck at the point of founding, followed by a period of rapid growth. We used ABC to test the posterior probabilities of these two hypotheses and to estimate parameters for the preferred scenario. All ABC analysis was conducted with the software package DIYABC (*ver. 1.0.4.46*) (Cornuet *et al.* 2008).

Prior distributions and summary statistics used in ABC

In order to specify informative priors we used information from case studies of infestations. Across 83 sampled infestations there was a range from eight to approximately 100,000 individuals (How & Lee 2010; Wang *et al.* 2010; Naylor 2012), with a geometric mean of 93 individuals. We therefore initially selected a prior range for the average effective size of an infestation at the time of sampling (Ne) bounded between 10 and 3,000 with a loguniform distribution. However, initial model runs suggested this was an overestimate so we reduced the prior range to a loguniform distribution bounded between 10 and 100 (Table 1). Based on observed female fecundity, infestations can reach over 3,000 individuals in as little as between two and six generations. Due to this rapid growth we fixed the bottleneck at one generation and selected a loguniform prior for time of founding (t2) bounded between two and 10 generations. Previous studies and our own data have suggested that infestations are started with small numbers of founders (Doggett *et al.* 2004; Saenz *et al.*

2012); we therefore set a uniform prior for this parameter (Nb) bounded between two and 14 individuals. This captures the range of possibilities from a single, once-mated founding female through a single, multiply-mated female to a larger group of individuals. Population growth was simulated as a stepwise increase from Nb to Ne at time t2-1. For parameters where data were not available, we used a two-step approach starting with wide, uninformative priors and then using narrower ranges that still captured whole posterior distributions (Table 1). As there is no prior knowledge on the mutation rate of short tandem repeats in bed bugs, we made the assumption that microsatellite loci followed a generalised stepwise mutation model (Estoup *et al.* 2002). This model is defined by two parameters: μ, the mean mutation rate across loci and P, the parameter of the geometric distribution describing the number of repeat changes per mutation event. As implemented in DIYABC, each locus has its own μ_i and P_i drawn from a gamma distribution of mean μ and P, respectively, with shape parameter set to k=2. All loci had a possible range of 40 contiguous allelic states and we used DIYABC's default values for mean μ and P prior ranges (Table 1). We expected global diversity to depend on the product of the source population size, Ns, and the mutation rate μ, because the vast majority of mutations must have occurred before the first modelled colonisation events. We did not expect to be able to estimate Ns and μ independently but this does not impact our ability to estimate the parameters describing recent colonisation and expansion events.

The type and number of summary statistics used in ABC is important to the analysis (Beaumont 2010) but there is still no clear consensus on determining the optimal set of statistics (Csilléry *et al.* 2010). We selected summary statistics that had been used

277 successfully in previous population genetic ABC studies, providing evidence that they contain
278 useful information, and that were available in DIYABC, a practical constraint (Cornuet *et al.*
279 2008; Lombaert *et al.* 2011; Dutech *et al.* 2012). These were mean number of alleles, mean
280 genic diversity (Nei 1978) and mean allele-size variance for each population and pairwise
281 population comparisons of F_{ST} (Weir & Cockerham 1984), giving a total of 39 single sample
282 statistics and 78 pairwise comparisons. Goodness-of-fit analyses were later used to check
283 the robustness of the results obtained using these summary statistics (see parameter
284 estimation and model checking).
285
286 *Simulation and model posterior probabilities.*
287 We simulated 1 x 10^6 genetic datasets for each of the two scenarios. As an initial
288 check that these scenarios could simulate datasets close to our observed data, we
289 performed a principal components analysis (PCA) on the summary statistics of the first
290 100,000 simulated datasets, and evaluated the position of our observed data. To estimate
291 the posterior probabilities of both scenarios a polychotomous weighted regression was
292 performed on the closest 1% of simulated data to the observed data (Cornuet *et al.* 2008).
293 Confidence in scenario choice was calculated by estimating false discovery rate as in Cornuet
294 *et al.* (2010) using 500 pseudo-observed datasets (POD), simulated under the scenario with
295 the lower posterior probability.
296
297 *Parameter estimation and model checking*
298 Assessing the goodness-of-fit (GoF) of the chosen model is a critical step in ABC
299 analyses (Gelman et al. 1995). We therefore ran posterior predictive simulations by creating

500 PODs from the posteriors of the chosen model. A PCA was then performed on summary statistics from (i) these 500 PODs, (ii) 500 datasets of each model randomly obtained from their priors and (iii) our observed data (Cornuet *et al.* 2010). As in Cornuet *et al.* (2010), we used a different set of summary statistics to perform the model checking compared to those used to calculate the posterior distributions of parameters to avoid overestimation of fit. We used Garza-Williamson's M (Garza and Williamson 2001) as a single sample statistic, and mean genic diversity and classification index (Rannala and Mountain 1997) between populations. The probability that the observed and the simulated data were significantly different was calculated by ranking the observed value of each test statistic against those obtained from the simulated data. *P*-values were corrected for multiple comparisons (Benjamini & Hochberg 1995).

To evaluate the posterior distribution of parameters, we performed a local linear regression on the closest 1% of logit-transformed simulated data. To assess confidence in our parameter estimates, 500 PODs were created using values drawn from our prior distribution. Using these datasets we calculated the relative median of the absolute error (RMAE) to estimate differences in the point estimates from the true values (Cornuet *et al.* 2010).

RESULTS

Characterising C. lectularius *infestations*

In total, 21 microsatellite markers were found to be polymorphic and were assembled into five ABI four-dye multiplex sets using the program Multiplex Manager

(Holleley & Geerts 2009) (Supporting Information, Table S2). To assess within-infestation genetic structure, 154 samples were genotyped from within five infestations at 19 of these loci. At the level of the whole infestation only LON2 showed a significant departure from HWE after Bonferroni correction. Overall, low genetic diversity was detected within infestations with a high number of monomorphic loci, low allelic richness, and low observed heterozygosity (Table 2). An F_{IS} value of -0.216 suggested an excess of heterozygotes within the AUS infestation (Table 2), which may be a signature of a recent bottleneck (Cornuet & Luikart 1996). Overall, estimates of F_{IS} were variable between infestations. This variability probably also reflects small numbers of founders. Note that the estimates are based on variable numbers of loci because of the occurrence of monomorphic loci within each infestation, consistent with low overall diversity, especially within BIR1, BIR2, LON1 and LON2. Within infestations, significant differentiation between refugia was only observed within LON2 (F_{ST} = 0.144, P=0.008; Table 2). Five out of 55 pairwise F_{ST} comparisons between LON2 refugia were significant, and no refuge was significantly differentiated from more than 27% of the other refugia (Table S3). In contrast, pairwise F_{ST} comparisons between infestations ranged from 0.492 to 0.834 (Table S4) and DAPC (Figure S2) showed very high levels of differentiation between infestations.

City-wide genetic diversity

In total, 63 individuals from 13 infestations across London were genotyped at 21 loci. Due to the low sample sizes per infestation (three to seven individuals) HWE could not be rejected. Within infestations, effective allele numbers ranged from 1.28 to 2.10 across all loci (Table 3) and were therefore comparable to more thoroughly sampled infestations (see

above and Booth *et al.* 2012; Saenz *et al.* 2012). Within infestations, mean kinship was high (0.566 ± 0.030) and across infestations F_{IS} was 0.084 (95% CI = -0.052 – 0.226). The global F_{ST} value across all 13 infestations showed significant population differentiation (F_{ST} = 0.592, SE = 0.026, p<0.001). However, we found no significant pattern of isolation by distance (Figure 2, slope = 0.013 ± 0.016). DAPC analysis further supported high differentiation between infestations with a defined genetic cluster for each infestation (Figure S3).

ABC analysis

A principal components analysis showed that our propagule pool model produced datasets that closely matched the observed data (Figure S4). The model comparison gave very strong support to scenario two. In fact, the closest 1% of datasets were all generated by scenario two. We therefore had to use 2% of simulated datasets to perform the logistic regression, which again gave almost complete support to scenario two (propagule pool model: posterior probability 0.9998, 95% Highest Posterior Density (HPD95): 0.9987, 1.0000) over scenario one (migrant pool model: 0.0002, HPD95: 0.0000, 0.0013). False discovery rate was low (5.2%) indicating the robustness of our model selection. Posterior predictive simulations showed that scenario two provides simulated datasets reasonably close to observed data (Figure S5). The deviation on the 2nd axis can be explained by the deviation of Garza-Williamson's M test statistics (the significant deviations are shown in Table S5).

With a posterior modal estimation of 3, the number of founders (Nb) starting each new infestation seems to be very limited (Median = 5.27, HPD90 = 2.09, 13.1 ; see Table 1). There was limited information regarding effective size of either current infestations or the source

population (Ne and Ns in Figure 3 and Table 1). The time since founding (t2) was low, suggesting that the sampled infestations were detected early (Median = 3.13, HPD90= 1.75, 9.36). Whilst the mutation rate parameter was not well estimated, additional simulations showed that adjusting the prior on this parameter did not have a strong effect on the other parameters of interest and resulted in an estimate of mutation rate within the range of the original prior (Table S6, Figure S6).

DISCUSSION

Using a combination of descriptive and ABC genetic analyses within a metapopulation framework, we have, for the first time, explicitly tested two competing models of bed bug colonisation dynamics. First, we found very low within-infestation diversity but despite this there is often rapid population growth. Second, ABC estimation favours a propagule pool model that, in combination with low observed diversity within infestations and very high differentiation between them, suggests that founders are related (Whitlock & McCauley 1990). This is also consistent with strong genetic homogeneity between refugia of the same dwelling, which is indicative of a single founding event per infestation. Low F_{ST} within infestations suggests that the infestation is the lowest level of population structure and not the refuge. Very high differentiation between infestations suggests limited connectivity and the lack of isolation by distance at the city scale fits our prediction that passive dispersal has weakened the relationship between genetic and

391 geographic distance (Johannesen & Lubin 1999; De Meester *et al.* 2002; Colson & Hughes
392 2004).
393
394 *Limitations and promise of our ABC approach*
395 Although theoretical studies have shown that the propagule and the migrant pool
396 models may produce somewhat similar differentiation patterns (Wade & McCauley 1988;
397 Pannell & Charlesworth 1999), our results clearly show that an ABC approach based on
398 multiple summary statistics can easily distinguish them. Using a high number of populations
399 with ABC means there may be a high dimensionality to the analysis, as it can inflate the
400 number of summary statistics used. This can be problematic as it often limits the ability to
401 simulate data close enough to the observed data, thus leading to inaccurate and potentially
402 biased scenario choices and parameter estimates (Beaumont *et al.* 2002; Blum *et al.* 2012).
403 In our case, the large number of summary statistics was primarily due to the use of pairwise
404 F_{ST} statistics. However, our PCA clearly shows that scenario two had the potential to simulate
405 data that were representative of our observed data (Figure S4: 90% of the variance
406 explained by both axes). The GoF analysis (Figure S5) further suggests that our choice of
407 summary statistics, whilst perhaps not optimal, was still good enough to make valid
408 inferences. It should also be noted that whilst our models are relatively simple they are
409 biologically relevant to this system in that they capture the key phases of colonisation,
410 growth and extinction. A more "realistic" model may have come at the cost of information
411 loss for our parameters of interest. In fact, we simulated a 3rd scenario in addition to the two
412 presented here. This scenario was similar to model two but with an additional bottleneck
413 after the initial divergence from the source population (i.e. a bottleneck between times t3

and t2). Despite this being more realistic, we lost information about our demographic parameters and thus we chose to disregard this model.

Scenario two, arguably the more biologically realistic of the two scenarios, was favoured and posterior predictive simulations gave good predictions of the observed data. The deviations observed for some test statistics between their predicted distributions and their observed values are likely a result of this being a simplified model. As noted above, we did not expect to be able to obtain robust estimates of Ns and μ, which together determine global population diversity. Adjusting the mutation parameter priors had little effect on the key parameters of interest (see Supplementary Information). Within-infestation diversity was expected to depend on Ne and Nb as well as the time parameters. The lack of information for Ne is likely due to the information loss caused by the strong bottlenecks and very fast subsequent population expansions within infestations. As there was a maximum of 10 generations between bottlenecks and sampling, there was not enough time for new mutations to arise, and so there was little constraint on the estimate of infestation size. Whilst care must be taken when using the ABC estimates for interpretation, it is important to note that when taken in combination with the genetic analysis our models produce biologically realistic estimates. This confirms the ability of ABC to make estimates of demographic parameters when using multiple populations and whilst working within a very short evolutionary time scale.

Bed bug infestation population dynamics

The very low levels of genetic diversity observed within infestations in this study, suggest low numbers of colonists. These patterns are reflected in other descriptive surveys of the genetic structure of bed bug populations (Booth *et al.* 2012; Saenz *et al.* 2012). High null allele frequency estimates and departures from Hardy-Weinberg equilibrium have been found in other genetic surveys of highly structured metapopulations (Johannesen & Lubin 1999; Massonnet *et al.* 2002; Kankare *et al.* 2004; Orsini *et al.* 2008). Despite several markers having high null allele frequency estimates, from our knowledge of bed bug life history in combination with previously published studies it is likely that these estimates are an artefact of the lower than expected heterozygosity resulting from founder events and inbreeding.

Due to living in aggregations and a high male mating rate, it is likely that most adult female bed bugs have been multiply mated (Stutt & Siva-Jothy 2001; Reinhardt *et al.* 2011). With low levels of genetic diversity a female that has been mated over 10 times may have only "effectively" mated once or twice, as she will likely only encounter related males. The ABC estimates are consistent with this. The asymmetry of the posterior distribution of Nb results in a wide HPD but there is a strong mode at Nb = 3. This modal value potentially represents a single female mated several times by genetically similar males; a version of the propagule pool model. This can explain the genetic data where high relatedness and kinship are detected within infestations. F_{IS} values have varied considerably across studies, but with few significant departures from zero (Booth *et al.* 2012; Saenz *et al.* 2012). This is likely a result of a low number of colonists with chance allele frequency differences between male and

460 female founders resulting in variation in the proportion of heterozygotes relative to Hardy-
461 Weinberg equilibrium. Despite low diversity, infestations can rapidly expand, suggesting
462 there are limited costs to inbreeding, and this is supported by preliminary experiments
463 performed by Johnson (1941). This may be due to the continuous purging of deleterious
464 alleles through repeated founder events (Hedrick 1994; Facon *et al.* 2011). Further work is
465 needed to examine the true cost of inbreeding in bed bugs.
466
467 Whilst generally consistent patterns of low diversity have been observed within sub-
468 populations, there is some variation in the genetic composition of infestations. For example,
469 some studies have reported higher genetic diversity within infestations, which would be
470 more consistent with the migrant pool model. Szalanski *et al.* (2008) reported one to six
471 different mitochondrial haplotypes within infestations in a survey of 11 different properties.
472 Booth *et al.* (2012) surveyed multiple apartments from residential blocks and found
473 evidence of sub-structuring within two of the three buildings. In contrast, the survey by
474 Saenz *et al.* (2012) showed that only one out of 21 properties had evidence of a multiple
475 introduction, therefore more closely following the predictions of the propagule pool model.
476 This variation suggests that the chance of multiple introductions is largely dependent on the
477 type of property (i.e. single residence or multi-dwelling), with a greater turnover of humans
478 increasing the likelihood of multiple founders. Due to the stochastic distribution of genetic
479 diversity generated by high levels of population turnover, large sample sizes are often
480 required to detect patterns of isolation by distance (Giles & Goudet 1997; Massonnet *et al.*
481 2002; Haag *et al.* 2005). As this study was not particularly designed to test for patterns of
482 isolation by distance we may have lacked the resolution to detect it and in fact some weak

patterns have been found in other studies (Saenz *et al.* 2012). Further work is still needed to determine the overall level of connectivity between sub-populations, and whether several large infestations are acting as sources for multiple patches or whether all infestations can be traced back to a larger mixed source population in an area where bed bug numbers have been consistently high.

Metapopulation dynamics

Due to discontinuous habitat, and small local population sizes, metapopulation dynamics should be fairly common in insects. However, classic metapopulations appear to be comparatively rare (Driscoll 2008; Driscoll *et al.* 2010), potentially because of low population turnover (Fronhofer *et al.* 2012). The close association between humans and bed bugs has led them to fulfil at least three of the four conditions required to be a classical metapopulation. Each separate infestation can be considered a discrete breeding patch, with bed bugs aggregating around a host food source, and infestations are likely to have similar effective population sizes. Pest control causes the independent local extinction of sub-populations and human facilitated dispersal gives the opportunity for recolonisation. However, reports of bed bug population expansion suggest that colonisation and extinction are not at equilibrium (Reinhardt & Siva-Jothy 2007). At present there is no precise information on the rate of population expansion, so the degree to which colonisation is outweighing extinction is not known.

505 We have demonstrated that bed bugs experience repeated founder events, with a severe
506 genetic bottleneck due to very low estimated numbers of founders. As predicted, the likely
507 common origin of colonists causes particularly strong differentiation (Whitlock & McCauley
508 1990). The relatively short life span of each infestation makes it unlikely that there will be an
509 introduction of novel alleles via gene flow, thus maintaining these levels of differentiation.
510 The low diversity within sub-populations may be further shaped by selective sweeps from a
511 strong insecticide selection pressure. Examples of species that can be considered
512 metapopulations have been shown to exist on a continuum, ranging from migrant pool
513 colonisers (Giles & Goudet 1997; Colson & Hughes 2004; Yang *et al.* 2008), to intermediates
514 (Whitlock 1992; Austin *et al.* 2010) and examples of high likelihood of a common origin of
515 founders (Ingvarsson *et al.* 1997). From our ABC estimates, alongside knowledge of the bed
516 bug life history, it is highly likely that in the majority of cases bed bug founders within an
517 infestation have a common origin. In combination these dynamics make bed bugs an
518 excellent model system to further investigate human impact on population structure and its
519 implication for diversity, as well as pest control.
520
521 *Implications for control*
522 Metapopulation frameworks have long been used to formulate efficient pest
523 management strategies (e.g. Levins, 1969; Cloarec 1999; Booth et al. 2011). Here, the
524 general low observed diversity within infestations suggests a single introduction to each
525 infestation. Type of dwelling, however, is likely to strongly influence the chances of having
526 multiple introductions. Hotels, apartment blocks and hospitals, which have the highest
527 turnover of visitors, are likely to be the most at risk from multiple introductions (Doggett &

Russell 2008). However, from the reported data so far it seems these cases may, for the moment at least, be in the minority and reported repeat infestations may be the result of failure to fully eradicate existing infestations (Boase 2008). The novel microsatellite markers described here, in conjunction with those described in Booth et al. (2012), could be used to determine whether a repeat infestation is a pest management failure or a recolonisation event by surveying properties before or immediately after treatment, and keeping specimens in case of a repeat infestation. This study also shows that due to the low diversity within and high differentiation between infestations, only a relatively small number of individuals is required to test for kinship between samples. Another factor to consider is that bed bugs are becoming rapidly resistant to insecticides. Infestations containing resistant individuals are more likely to avoid extinction, prolonging their time as a source. Resistant alleles would be quickly selected for and become fixed in populations. These markers, in combination with candidate resistance loci, can now be used to track resistance alleles moving through metapopulations.

Acknowledgements

We would like to thank Richard Naylor and Oliver Otti for support throughout the project. The laboratory work was performed at the NERC Biomolecular Analysis Facility at the University of Sheffield, UK (microsatellite cloning and genotyping) and University of Edinburgh (Sanger sequencing) supported by the Natural Environment Research Council (NERC), UK. We thank Deborah Dawson and Terry Burke for advice and support, Jill Lovell and Andy Gillies for sequencing, Andy Krupa for genotyping advice and all those who

contributed samples, particularly David Cain. Finally we thank Arnaud Estoup for advice and support with DIYABC. TF was funded by a NERC postgraduate studentship and support from the Royal Entomological Society. KR is supported by an advanced postdoctoral fellowship of the Volkswagen Foundation.

REFERENCES

Austin A, Ovaskainen O, Hanski I (2010) Size and genetic composition of the colonizing propagules in a butterfly metapopulation. *Oikos*, **120**, 1357–1365.

Beaumont MA (2010) Approximate Bayesian Computation in Evolution and Ecology. *Annual Review of Ecology, Evolution, and Systematics*, **41**, 379–406.

Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics*, **162**, 2025–2035.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289–300.

Blum MGB, Nunes MA, Prangle D, Sisson SA (2012) A comparative review of dimension reduction methods in approximate Bayesian computation. *arXiv preprint arXiv:1202.3819*.

Boase CJ (2001) Bedbugs-back from the brink. *Pesticide Outlook*, **12**, 159–162.

Boase CJ (2008) Bed bugs (Hemiptera: Cimicidae): an evidence-based analysis of the current situation. *Proceedings of the 6th International Conference on Urban Pests, OOK-Press Kft., Hungary*, 7–14.

Booth W, Satangelo RG, Vargo E, Mukha DV, Schal C (2011) Population Genetic Structure in German Cockroaches (Blattella Germanica): Differentiated Islands in an Agricultural Landscape. Journal of Heredity, 102, 2, 175-183.

Booth W, Saenz VL, Santangelo RG *et al.* (2012) Molecular markers reveal infestation dynamics of the bed bug (Hemiptera: Cimicidae) within apartment buildings. *Journal of Medical Entomology*, **49**, 535–546.

Churcher TS, Schwab AE, Prichard RK, Basáñez M-G (2008) An Analysis of Genetic Diversity and Inbreeding in Wuchereria bancrofti: Implications for the Spread and Detection of Drug Resistance (J McCarthy, Ed,). *PLoS neglected tropical diseases*, **2**, e211.

Collins FH, Kamau L, Ranson HA, Vulule JM (2000) Molecular entomology and prospects for malaria control. *Bulletin of the World Health Organization*, **78**, 1412–1423.

Colson I, Hughes RN (2004) Rapid recovery of genetic diversity of dogwhelk (*Nucella lapillus* L.) populations after local extinction and recolonization contradicts predictions from life-history characteristics. *Molecular Ecology*, **13**, 2223–2233.

Cornuet J-M, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.

590 Cornuet J-M, Ravignie V, Estoup A (2010) Inference on population history and model
591 checking using DNA sequence and microsatellite data with the software DIYABC (v1.0).
592 *BMC Bioinformatics*, **11**, 401.
593 Cornuet J-M, Santos F, Beaumont MA *et al.* (2008) Inferring population history with DIY ABC:
594 a user-friendly approach to approximate Bayesian computation. *Bioinformatics*, **24**,
595 2713–2719.
596 Cloarec A, Rivault C, Cariou ML. (1999) Genetic population structure of the German
597 cockroach, *Blattella germanica*: absence of geographical variation. Entomol Exp Appl.,
598 **92**, 311–319.
599 Csilléry K, Cilas C, Blum MGB, Gaggiotti OE, Francois O (2010) Approximate Bayesian
600 Computation (ABC) in practice. *Trends in Ecology & Evolution*, 1–9.
601 De Meester L, Gómez A, Okamura B, Schwenk K (2002) The Monopolization Hypothesis and
602 the dispersal–gene flow paradox in aquatic organisms. *Acta Oecologica*, **23**, 121–135.
603 Dieringer D, Schlötterer C (2003) Microsatellite analyser (MSA): a platform independent
604 analysis tool for large microsatellite data sets. *Molecular Ecology Notes*, **3**, 167–169.
605 Doggett S, Russell R (2008) The resurgence of bed bugs, Cimex spp.(Hemiptera: Cimicidae) in
606 Australia. *Proceedings of the 6th International Conference on Urban Pests, OOK-Press*
607 *Kft., Hungary*, 407–425.
608 Doggett S, Geary M, Russell R (2004) The resurgence of bed bugs in Australia: with notes on
609 their ecology and control. *Environmental Health -Australian Institute of Environmental*
610 *Health*, **4**, 30–38.
611 Driscoll DA (2008) The frequency of metapopulations, metacommunities and nestedness in a
612 fragmented landscape. *Oikos*, **117**, 297–309.
613 Driscoll DA, Kirkpatrick JB, McQuillan PB, Bonham KJ (2010) Classic metapopulations are rare
614 among common beetle species from a naturally fragmented landscape. *Journal of*
615 *Animal Ecology*, **79**, 294–303.
616 Dutech C, Barrès B, Bridier J *et al.* (2012) The chestnut blight fungus world tour: successive
617 introduction events from diverse origins in an invasive plant fungal pathogen. *Molecular*
618 *Ecology*, **21**, 3931–3946.
619 Estoup A, Beaumont MA, Sennedot F, Moritz C, Cornuet J-M (2004) Genetic analysis of
620 complex demographic scenarios: spatially expanding populations of the cane toad, Bufo
621 marinus. *Evolution*, **58**, 2021–2036.
622 Estoup A, Jarne P, Cornuet J-M (2002) Homoplasy and mutation model at microsatellite loci
623 and their consequences for population genetics analysis. *Molecular Ecology*, **11**, 1591–
624 1604.
625 Facon B, Hufbauer RA, Tayeh A *et al.* (2011) Inbreeding Depression Is Purged in the Invasive
626 Insect Harmonia axyridis. *Current Biology*, **21**, 424–427.
627 Frankham R, Ballou JD, Briscoe DA (2002) *Introduction to Conservation Genetics*. Cambridge
628 University Press.
629 Fronhofer EA, Kubisch A, Hilker FM, Hovestadt T, Poethke HJ (2012) Why are
630 metapopulations so rare? *Ecology*, **93**, 1967–1978.
631 Gaggiotti OE, Brooks SP, Amos W, Harwood J (2004) Combining demographic, environmental
632 and genetic data to test hypotheses about colonization events in metapopulations.
633 *Molecular Ecology*, **13**, 811–825.
634 Garza J, Williamson E (2001) Detection of reduction in population size using data from

microsatellite loci. *Molecular Ecology*, **10**, 305–318.
Gelman, A, Carlin JB, Stern HS and Rubin DB (1995) *Bayesian Data Analysis*. Chapman et Hall, London.
Giles BE, Goudet J (1997) Genetic differentiation in *Silene dioica* metapopulations: estimation of spatiotemporal effects in a successional plant species. *The American Naturalist*, 507–526.
Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity*, **86**, 485–486.
Grapputo A, Boman S, Lindstroem L, Lyytinen A, Mappes J (2005) The voyage of an invasive species across continents: genetic diversity of North American and European Colorado potato beetle populations. *Molecular Ecology*, **14**, 4207–4219.
Haag CR, Riek M, Hottinger JW, Pajunen VI, Ebert D (2005) Founder events as determinants of within-island and among-island genetic structure of Daphnia metapopulations. *Heredity*, **96**, 150–158.
Hanski I (1999) *Metapopulation Ecology*. Oxford University Press.
Hardy O, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.
Hastings A, Harrison S (1994) Metapopulation dynamics and genetics. *Annual Review of Ecology and Systematics*, 167–188.
Hedrick PW (1994) Purging Inbreeding Depression and the Probability of Extinction - Full-Sib Mating. *Heredity*, **73**, 363–372.
Holleley C, Geerts P (2009) Multiplex Manager 1.0: a cross-platform computer program that plans and optimizes multiplex PCR. *BioTechniques*, **46**, 511–517.
How Y, Lee C (2010) Survey of Bed Bugs in Infested Premises in Malaysia and Singapore. *Journal of Vector Ecology*, **35**, 89–94.
Ingvarsson PK (1998) Kin-structured colonization in *Phalacrus substriatus*. *Heredity*, **80**, 456–463.
Ingvarsson PK, Olsson K (1997) Hierarchical genetic structure and effective population sizes in *Phalacrus substriatus*. *Heredity*, **79**, 153–161.
Ingvarsson PK, Olsson K, Ericson L (1997) Extinction-recolonization dynamics in the mycophagous beetle *Phalacrus substriatus*. *Evolution*, 187–195.
Johannesen J, Lubin Y (1999) Group founding and breeding structure in the subsocial spider *Stegodyphus lineatus* (Eresidae). *Heredity*, **82**(6), 677–686.
Johnson CG (1941) The ecology of the bed-bug, Cimex lectularius L., in Britain: Report on Research, 1935-40. *The Journal of Hygiene*, **41**, 345–461.
Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403–1405.
Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, **11**, 94.
Kalinowski, ST, Taper, ML & Marshall, TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology,* **16**, 1099-1006.
Kankare M, Nouhuys S, Gaggiotti O, Hanski I (2004) Metapopulation genetic structure of two coexisting parasitoids of the Glanville fritillary butterfly. *Oecologia*, **143**, 77–84.

Kilpinen O, Jensen KMV, Kristensen M (2008) Bed bug problems in Denmark, with a European perspective. *Proceedings of the Sixth International Conference on Urban Pests*, 13–16.

Lawson Handley LJ, Estoup A, Evans DM *et al.* (2011) Ecological genetics of invasive alien species. *BioControl*, **56**, 409–428.

Levins R (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America,* **15**, 3, 237-240

Loiselle B, Sork V, Nason J, Graham C (1995) Spatial Genetic-Structure of a Tropical Understory Shrub, Psychotria Officinalis (Rubiaceae). *American Journal of Botany*, **82**, 1420–1425.

Lombaert E, Guillemaud T, Thomas CE *et al.* (2011) Inferring the origin of populations introduced from a genetically structured native range by approximate Bayesian computation: case study of the invasive ladybird *Harmonia axyridis*. *Molecular Ecology*, **20**, 4654–4670.

Massonnet B, Simon J-C, Weisser WW (2002) Metapopulation structure of the specialized herbivore *Macrosiphoniella tanacetaria* (Homoptera, Aphididae). *Molecular Ecology*, **11**, 2511–2521.

Naylor R (2012) Ecology and dispersal of the bedbug. *PhD Thesis, University of Sheffield.*

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.

Nicholls J, Double M, Rowell D, Magrath R (2000) The evolution of cooperative and pair breeding in thornbills Acanthiza (Pardalotidae). *Journal of Avian Biology*, **31**, 165–176.

Niggemann M, Jetzkowitz J, Brunzel S, Wichmann MC, Bialozyt R (2009) Distribution patterns of plants explained by human movement behaviour. *Ecological Modelling*, **220**, 1339–1346.

Orsini L, Corander J, Alasentie A, Hanski I (2008) Genetic spatial structure in a butterfly metapopulation correlates better with past than present demographic structure. *Molecular Ecology*, **17**, 2629–2642.

Pannell JR, Charlesworth B (1999) Neutral genetic diversity in a metapopulation with recurrent local extinction and recolonization. *Evolution*, **53**, 664–676.

Potter M, Romero A, Haynes K (2008) Battling bed bugs in the USA. *Proceedings of the 6th International Conference on Urban Pests, OOK-Press Kft., Hungary*, 400–406.

R Core Team (2012) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 9197–9201.

Ray C (2001) Maintaining genetic diversity despite local extinctions: effects of population scale. *Biological Conservation*, **100**, 3–14.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of heredity*, **86**, 248–249.

Reinhardt K, Siva-Jothy MT (2007) Biology of the Bed Bugs (Cimicidae). *Annual Review of Entomology*, **52**, 351–374.

Reinhardt K, Isaac D, Naylor R (2010) Estimating the feeding rate of the bedbug *Cimex lectularius* in an infested room: an inexpensive method and a case study. *Medical and Veterinary Entomology*, **24**, 46–54.

Reinhardt K, Naylor R, Siva-Jothy MT (2011) Male Mating Rate Is Constrained by Seminal Fluid Availability in Bedbugs, *Cimex lectularius* (B Fenton, Ed,). *PLoS ONE*, **6**, e22082.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

Richards L, Boase CJ, Gezan S, Cameron M (2009) Are bed bug infestations on the increase within Greater London? *Journal of Environmental Health Research*, **9**, 17–24.

Rinkevich FD, Hamm RL, Geden CJ, Scott JG (2007) Dynamics of insecticide resistance alleles in house fly populations from New York and Florida. *Insect Biochemistry and Molecular Biology*, **37**, 550–558.

Romero A, Potter M, Potter D, Haynes K (2007) Insecticide resistance in the bed bug: a factor in the pest's sudden resurgence? *Journal of Medical Entomology*, **44**, 175–178.

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Rousset F (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.

Saenz VL, Booth W, Schal C, Vargo EL (2012) Genetic analysis of bed bug populations reveals small propagule size within individual infestations but high genetic diversity across infestations from the eastern United States. *Journal of Medical Entomology*, **49**, 865–875.

Slatkin M (1977) Gene flow and genetic drift in a species subject to frequent local extinctions. *Theoretical Population Biology*, **12**, 253-262

Stutt A, Siva-Jothy M (2001) Traumatic insemination and sexual conflict in the bed bug *Cimex lectularius*. *Proceedings of the National Academy of Sciences*, **98**, 5683.

Szalanski A, Austin J, McKern J, Steelman C, Gold R (2008) Mitochondrial and ribosomal internal transcribed spacer 1 diversity of *Cimex lectularius* (Hemiptera: Cimicidae). *Journal of Medical Entomology*, **45**, 229–236.

Tatem A, Hay S, Rogers D (2006) Global traffic and disease vector dispersal. *Proceedings of the National Academy of Sciences*, **103**, 6242-6247.

Torimaru T, Tani N, Tsumura Y, Nishimura N, Tomaru N (2007) Effects of kin-structured seed dispersal on the genetic structure of the clonal dioecious shrub *Ilex leucoclada*. *Evolution*, **61**, 1289–1300.

Wade MJ, McCauley DE (1988) Extinction and Recolonization - Their Effects on the Genetic Differentiation of Local-Populations. *Evolution*, **42**, 995–1005.

Walser B, Haag CR (2012) Strong intraspecific variation in genetic diversity and genetic differentiation in *Daphnia magna*: the effects of population turnover and population size. *Molecular Ecology*, **21**, 851–861.

Wang C, Saltzmann K, Chin E, Bennett GW, Gibb T (2010) Characteristics of *Cimex lectularius* (Hemiptera: Cimicidae), Infestation and Dispersal in a High-Rise Apartment Building. *Journal of Economic Entomology*, **103**, 172–177.

Weir B, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population-Structure. *Evolution*, **38**, 1358–1370.

Whitlock MC (1992) Nonequilibrium Population-Structure in Forked Fungus Beetles - Extinction, Colonization, and the Genetic Variance Among Populations. *The American*

Naturalist, **139**, 952–970.
Whitlock MC, McCauley DE (1990) Some Population Genetic Consequences of Colony Formation and Extinction - Genetic Correlations Within Founding Groups. *Evolution*, **44**, 1717–1724.
Yakob L, Kiss IZ, Bonsall MB (2008) A network approach to modeling population aggregation and genetic control of pest insects. **74**, 324–331.
Yang S, Bishop JG, Webster MS (2008) Colonization genetics of an animal-dispersed plant (*Vaccinium membranaceum*) at Mount St Helens, Washington. *Molecular Ecology*, **17**, 731–740.

Figure 1 Demographic scenarios for ABC. *Scenario one: Migrant pool model of colonisation*: All populations originate from a single hypothetic source population (Ns), which represents the metapopulation as a whole. At time t2 these populations diverge signifying the founding of new infestations, which includes a severe bottleneck (Nb). The populations subsequently grow and after a bottleneck of one generation reach effective population size (Ne). It is at this size at which the infestations are sampled. *Scenario two: Propagule pool model of colonisation*: In this scenario founders diverge from Ns at t3 and maintain a population size of Ne until t2 where there is a founding event and a severe bottleneck (Nb). The sub-population then grows to a size of Ne after a bottleneck of 1 generation before being sampled. In both scenarios only infestations a and b are shown, but models incorporate 13 sampled infestations, represented by the dotted line.

Figure 2 Kinship plotted against distance with standard error bars. The first point represents within-infestation kinship, the following 10 points represent geographic distance, which was broken down into 10 distance intervals by SPAGeDi.

Figure 3 Prior (Grey) and Posterior (Black) distributions of parameters obtained under the better-supported model (scenario two). The x-axis shows the range of parameter values, and the y-axis the probability density.

Data accessibility

The 329 microsatellite sequences isolated during the development of the genomic library including those of the 21 loci characterised in this study have been submitted to the EMBL database (HF969864–HF970194) and microsatellite genotyping data are available in DRYAD, doi:XXXXXXXXX.

Author contributions

TF, KR and RKB designed the study. TF and GH performed the experiments. TF and LD performed the analysis. TF, LD, KR and RKB wrote the paper.

Table 1 Details of prior and posterior distributions of model parameters. Parameters constrained such that Ns>Ne>Nb.

Parameters	Prior range	Mean	Median	Mode	HPD90 low	HPD90 high	RMAE
Ns	Loguniform [100 - 50000]	6320	4950	2460	1520	15200	0.416
Ne	Loguniform [10 - 100]	33.5	32.6	35.6	12.7	57.1	0.258
Nb	Uniform [2 - 14]	6.21	5.27	3.00	2.09	13.1	0.345
t2	Loguniform [2 - 10]	3.97	3.13	2.00	1.75	9.36	0.456
t3	Uniform [11 - 100]	52.2	49.6	26.8	18.3	94.1	0.266
Mean μ	Loguniform [10^{-4} - 10^{-3}]	3.02 x 10^{-4}	2.22 x 10^{-4}	1.00 x 10^{-4}	1.07 x 10^{-4}	7.86 x 10^{-4}	0.382
Mean P	Uniform [0.1 - 0.3]	0.115	0.103	0.100	0.100	0.167	0.243

RMAE = Relative Median of the Absolute Error.
RMAE computed using 500 pseudo-observed datasets taking the median of posterior distribution as point estimates

Table 2 Genetic diversity and structure within five *C. lectularius* infestations for which multiple refugia were sampled. Total is the value obtained when individual for all localities were pooled together. Heterozygosity and F-statistics were calculated within and among *C. lectularius* infestations at 19 loci. Significance of F_{ST} values was calculated after 10,000 permutations.

Infestation	n	Allele Range	Allelic Richness[a] (± SE, n=17)	L_M	He	Ho	F_{IS} (95% CI)	F_{ST}
AUS	41	1-4	1.71 (0.15)	3	0.280	0.334	-0.216 (-0.349, -0.035)	0.017^{NS}
BIR1	8	1-3	1.61 (0.12)	7	0.173	0.054	0.727 (0.476, 0.926)	-0.184^{NS}
BIR2	9	1-5	2.10 (0.24)	6	0.380	0.337	0.141 (-0.104, 0.394)	-0.044^{NS}
LON1	52	1-3	1.19 (0.09)	13	0.075	0.084	-0.135 (-0.232, -0.013)	0.010^{NS}
LON2	46	1-3	1.62 (0.16)	8	0.250	0.163	0.219 (0.079, 0.475)	0.144^{**}
Total	156	1-5	2.98 (0.15)	7.4	0.566	0.194	0.052 (-0.072, 0.198)	0.709^{***}

n = Number of individuals, [a]Samples standardised to size of the smallest number of individuals for a locus
Two loci were omited because they failed to amplify for any individual in one infestation
L_M = Number of monomorphic loci, H_E = Expected heterozygosity, H_O = Observed heterozygosity
^{NS}non-significant (p>0.05), **significant (p<0.01), ***highly significant (p<0.001)

838 **Table 3** Indices of genetic diversity of 13 infestations from across London, UK. Total is the
839 value obtained when individual for all localities were pooled together

Infestation	Co-ordinates Lat	Long	n	E_A	Allelic Richness[a] (± SE, n=20)	H_E	H_O
a	51.4924	-0.2294	3	1.35	1.41 (0.12)	0.209	0.270
b	51.4693	-0.1138	3	1.62	1.65 (0.11)	0.384	0.333
c	51.5293	-0.0218	3	1.28	1.37 (0.10)	0.197	0.270
d	51.4924	-0.1674	7	1.37	1.41 (0.10)	0.223	0.211
e	51.5851	-0.2602	4	1.28	1.37 (0.10)	0.189	0.163
f	51.5333	-0.1681	7	1.54	1.59 (0.11)	0.303	0.213
g	51.4491	-0.1215	5	2.10	2.07 (0.13)	0.510	0.279
h	51.6905	-0.0338	4	1.51	1.58 (0.12)	0.287	0.298
i	51.5840	-0.1171	5	1.62	1.67 (0.09)	0.370	0.412
j	51.4799	-0.0296	4	1.49	1.55 (0.10)	0.308	0.345
k	51.3843	-0.4207	6	1.41	1.41 (0.10)	0.242	0.216
l	51.5113	-0.2679	6	1.42	1.43 (0.11)	0.228	0.213
m	51.5561	-0.1739	6	1.46	1.56 (0.11)	0.275	0.317
Total			63	3.26	2.53 (0.06)	0.680	0.266

[a]Samples standardised to size of the smallest number of individuals for a locus
One locus was omitted as for one infestation no individuals amplified for that locus
E_A = Effective number of alleles n = Number of individuals
H_E = Expected heterozygosity H_O = Observed heterozygosity

840

841

842

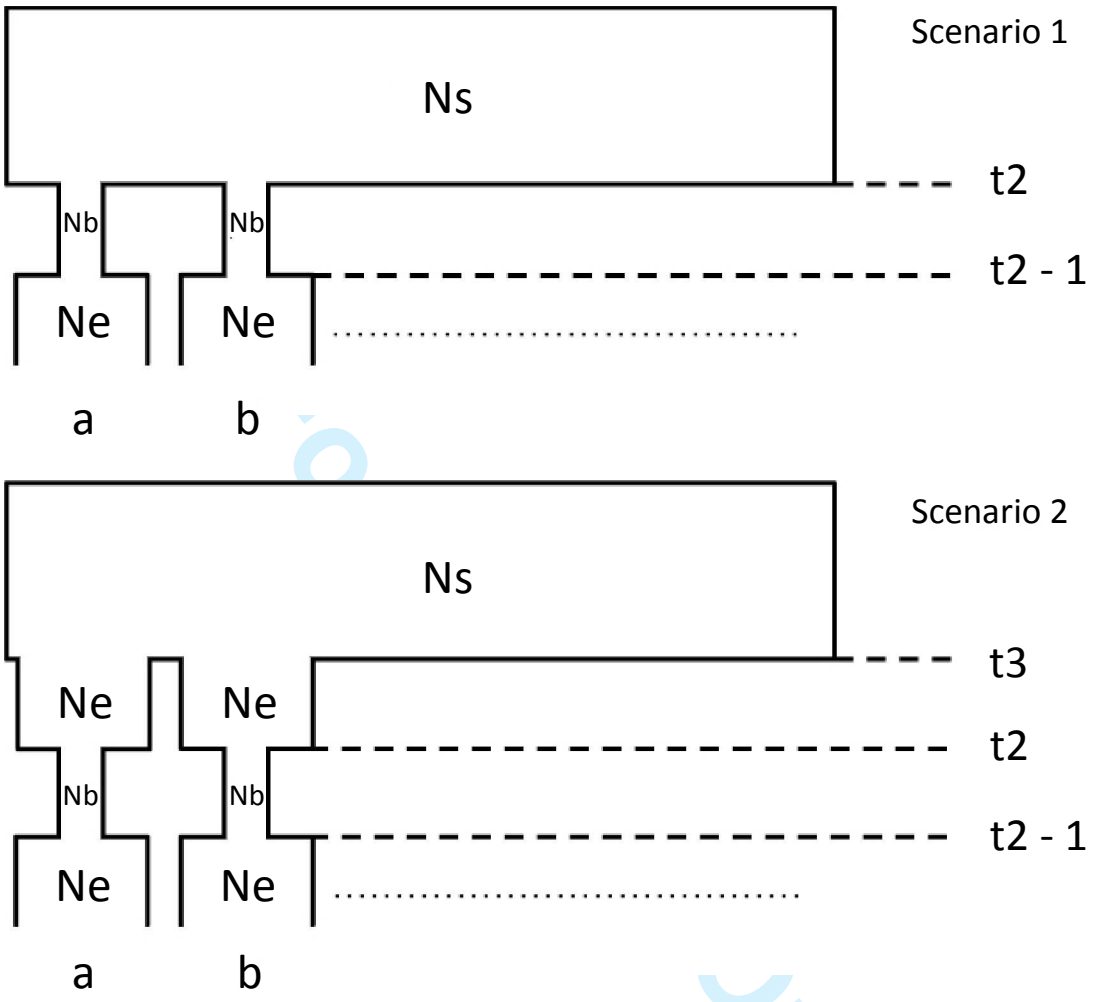
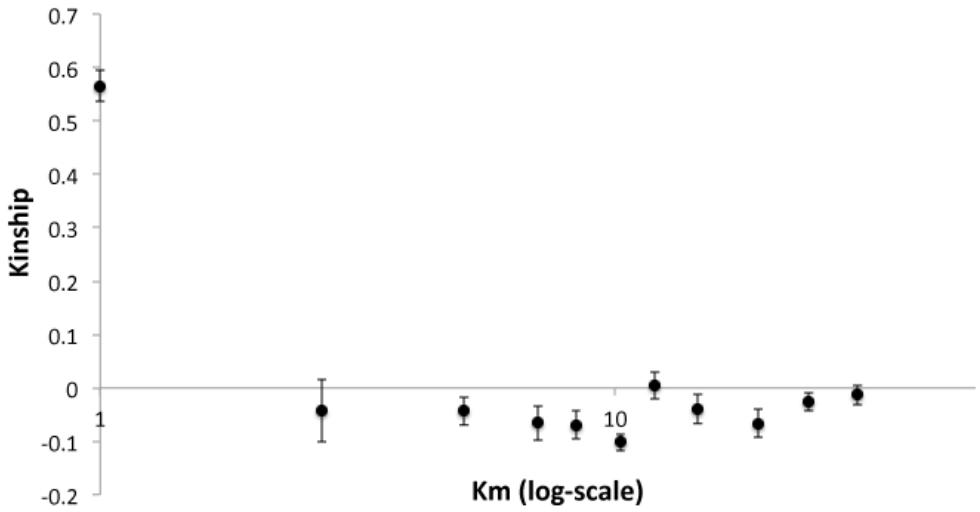

Scenario 1
Ns
t2
Nb
Nb
t2 - 1
Ne
Ne
a
b
Scenario 2
Ns
t3
Ne
Ne
t2
Nb
Nb
t2 - 1
Ne
Ne
a
b

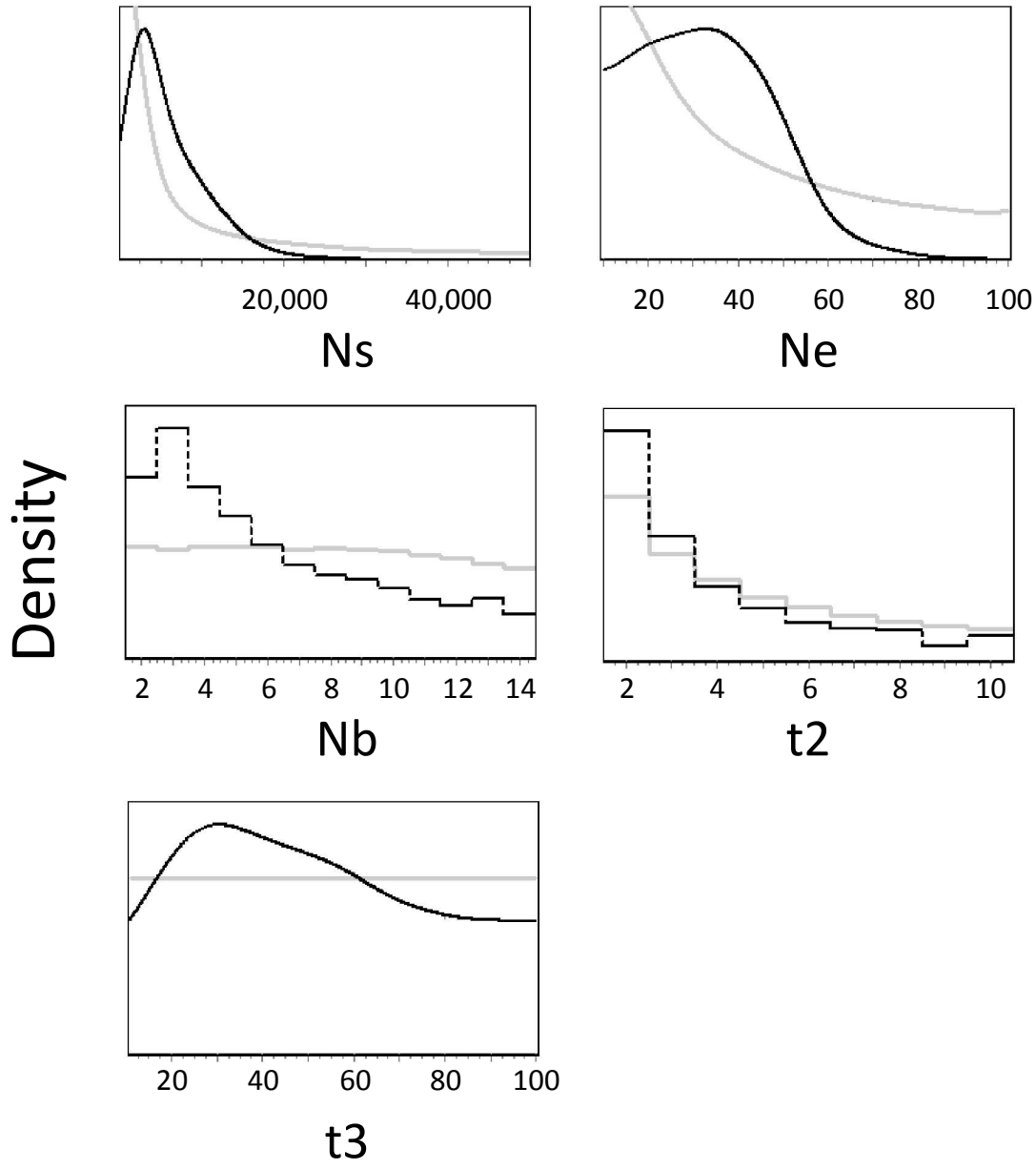
Figure 1

855



856

857 **Figure 2**



858

859 **Figure 3**