MOLECULAR ECOLOGY

Molecular Ecology (2015) **24**, 296–309 doi: 10.1111/mec.13025

Life-stage differences in spatial genetic structure in an irruptive forest insect: implications for dispersal and spatial synchrony

PATRICK M. A. JAMES,*[1] BARRY COOKE,† BRYAN M. T. BRUNET,‡ LISA M. LUMLEY,‡§ FELIX A. H. SPERLING,‡ MARIE-JOSÉE FORTIN,¶ VANESSA S. QUINN** and BRIAN R. STURTEVANT††[1]

*Département de Sciences Biologiques, CP 6128 Succursale Centre-Ville, Université de Montréal, Pavillon Marie-Victorin, Montréal, QC, Canada, H3C 3J7, †Canadian Forest Service, Northern Forestry Centre, 5320 122 Street Northwest, Edmonton, AB, Canada, T6H 3S5, ‡Department of Biological Sciences, University of Alberta, CW 405, Biological Sciences Bldg., Edmonton, AB, Canada, T6G 2E9, §Royal Alberta Museum, 12845 102 Avenue, Edmonton, AB, Canada, T5N 0M6, ¶Department of Ecology and Evolutionary Biology, University of Toronto, 25 Harbord St., Toronto, ON, Canada, M5S 3G5, **Purdue University – North Central, Westville, IN 46391, USA, ††Institute for Applied Ecosystem Studies, Northern Research Station, USDA Forest Service, 5985 Hwy K, Rhinelander, WI 54501, USA

Abstract

Dispersal determines the flux of individuals, energy and information and is therefore a key determinant of ecological and evolutionary dynamics. Yet, it remains difficult to quantify its importance relative to other factors. This is particularly true in cyclic populations in which demography, drift and dispersal contribute to spatio-temporal variability in genetic structure. Improved understanding of how dispersal influences spatial genetic structure is needed to disentangle the multiple processes that give rise to spatial synchrony in irruptive species. In this study, we examined spatial genetic structure in an economically important irruptive forest insect, the spruce budworm (*Choristoneura fumiferana*) to better characterize how dispersal, demography and ecological context interact to influence spatial synchrony in a localized outbreak. We characterized spatial variation in microsatellite allele frequencies using 231 individuals and seven geographic locations. We show that (i) gene flow among populations is likely very high ($F_{st} \approx 0$); (ii) despite an overall low level of genetic structure, important differences exist between adult (moth) and juvenile (larvae) life stages; and (iii) the localized outbreak is the likely source of moths captured elsewhere in our study area. This study demonstrates the potential of using molecular methods to distinguish residents from migrants and for understanding how dispersal contributes to spatial synchronization. In irruptive populations, the strength of genetic structure depends on the timing of data collection (e.g. trough vs. peak), location and dispersal. Taking into account this ecological context allows us to make more general characterizations of how dispersal can affect spatial synchrony in irruptive populations.

Keywords: cyclic populations, insect outbreaks, microsatellites, multivariate analysis, population genetics, spruce budworm

Received 5 August 2014; revision received 23 November 2014; accepted 26 November 2014

Introduction

Dispersal is one of the most important processes in animal ecology and is a primary determinant of population dynamics, spatial genetic structure and evolution. Dispersal also plays an important, yet not completely understood role in the spatial synchrony of population dynamics in cyclic species (Ehrich *et al.* 2001; Krebs *et al.* 2002; Peltonen *et al.* 2002; Schwartz *et al.* 2002; Franklin *et al.* 2014; Norén & Angerbjörn 2014). With sufficient recapture rates, traditional field methods can infer population dispersal rates but

Correspondence: Patrick M. A. James, Fax: 514-343-2293; E-mail: patrick.ma.james@umontreal.ca and Brian Sturtevant, Fax: 715-362-1166; E-mail: bsturtevant@fs.fed.us

[1]These authors contributed equally to this work.

cannot easily define the sources and destinations of dispersers nor mechanistically describe the spatial extent of synchrony. Alternatively, measures of spatial genetic structure can be used to estimate the frequency, intensity and spatial scale of effective dispersal (Broquet & Petit 2009). The interplay between multiple microevolutionary and demographic processes in cyclic populations (e.g. population irruption followed by collapse at differing spatial and temporal scales) makes it difficult to produce generalizations on what the spatial genetic structure of a cyclic system will look like over time.

Cyclic population dynamics and their genetic consequences have been an area of interest in ecology for many years (Elton 1924). Recent synthesis has demonstrated that dispersal is essential to understanding the feedback between population cycles and the spatial genetic variation (Norén & Angerbjörn 2014). However, the vast majority of studies investigating these relationships have focused on a limited range in taxa—most notably high-latitude mammals such as arvicolines (i.e. lemmings, voles, etc.) (Norén & Angerbjörn 2014). Relatively little is known about the influence of cyclic dynamics on the genetic structure in other taxa, such as insects, despite the prevalence of high amplitude cyclic behaviour in several species of economic importance in agro-forestry such as Lepidoptera in high-latitude forest systems (Myers & Cory 2013).

Particularly, vexing is the role of dispersal in the eastern spruce budworm (*Choristoneura fumiferana*) system. This system is particularly well suited for examining the question of synchronized population irruptions because of (i) its relatively long period between successive outbreaks (*c.* 35 years; Jardon *et al.* 2003) that we expect will result in notable genetic differentiation among populations between irruptions (Myers & Cory 2013); (ii) the large spatial scale of population synchronization (Peltonen *et al.* 2002); (iii) previously described dispersal capacity (Greenbank *et al.* 1980); and (iv) multiple focal outbreak locations that appear to coalesce over the development of the outbreak (Royama 1984; Box 1). Following a continental-scale analysis using isozymes, it was previously concluded there was little to no limitation to dispersal across much of its range (Harvey 1996). However, we suggest that comparing samples taken from different ecological (i.e. outbreak) contexts across the geographic range of a species may confound efforts to understand the demographic processes that lead to population genetic structure in irruptive systems. To date, there have been no additional published studies of eastern spruce budworm population genetic structure. Examination of the genetic structure of a spruce budworm outbreak within a specific ecological context has the potential to further elucidate the role of dispersal in synchronized irruptive systems.

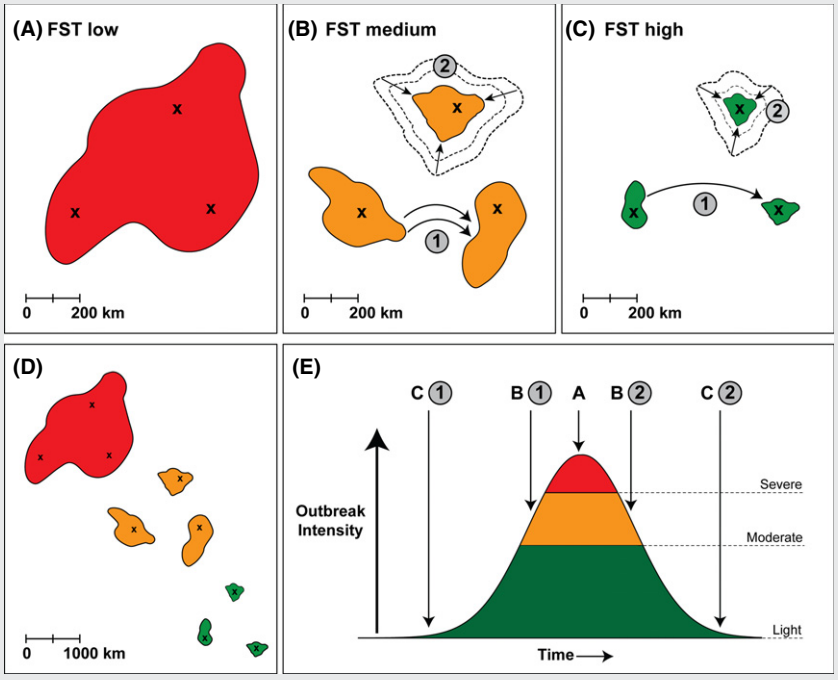
Budworm population dynamics are both cyclic and irruptive in that cycles result in increases over several orders of magnitude and have significant economic consequences (i.e. an 'outbreak'). These populations exhibit both local epicentric growth (Greenbank *et al.* 1980; Hardy *et al.* 1983) and regional-scale synchronization (up to 500 000 km^2) during outbreaks (Royama 1984). Dispersal from areas of high to low abundance can switch population dynamics from an endemic to an epidemic state, presumably by helping populations overcome local Allee effects driven by a low mate finding success at low densities (Régnière *et al.* 2013). Refining our understanding of dispersal and its genetic consequences in irruptive populations is thus of critical importance to improve our capacity to monitor, predict and manage outbreaks.

The spruce budworm is a univoltine native moth that defoliates balsam fir (*Abies balsamea*) and spruce species (*Picea* spp.) in the boreal and sub-boreal forests of North America (Royama 1984). Periodic outbreaks cause extensive forest mortality, reduce tree growth and influence timber and fibre production. Females lay eggs on fir and spruce in late summer that hatch in *c.* 10 days. First-instar larvae disperse a short distance to hibernacula and moult into second-instar larvae and enter diapause. Larvae emerge in May and begin feeding on foliage. Pupation occurs in mid-June, and moths eclose in *c.* 10 days, completing the cycle. Moths can easily disperse 20 km with a maximum recorded dispersal distance of 450 km (Greenbank *et al.* 1980). Long-distance dispersal is episodic and depends on atmospheric processes (Sturtevant *et al.* 2013), local population density (Royama 1984), the condition of host trees (Régnière & Nealis 2007) and outbreak progression (Royama 1984).

The spatial extent of outbreaks (i.e. spatial synchrony) combined with the degree of damage caused (i.e. directly correlated with population amplitude) determines the degree to which humans define an organism as a 'pest' (Liebhold *et al.* 2012). Three processes synchronize the dynamics of spatially disjunct populations (Bjørnstad *et al.* 1999; Liebhold *et al.* 2004): (i) the Moran effect in which spatially autocorrelated climate can synchronize population dynamics over broad spatiotemporal scales (Moran 1953); (ii) spatially synchronized populations of natural enemies that induce cycling in their host via density-dependent regulation (Bjørnstad & Bascompte 2001; Tobin & Bjørnstad 2003); and (iii) dispersal, through which independently cycling populations are linked (Williams & Liebhold 2000; Peltonen *et al.* 2002; Tobin & Bjørnstad 2005). Imperfect

Box 1.

Spatial genetic structure is influenced by ecological context. Genetic structure in cyclic and irruptive populations will depend on the timing of data collection (peak or trough), context-dependent dispersal in the species, as well as the amplitude and frequency of population oscillations (Norén & Angerbjörn 2014). Panels A, B and C show three population structure scenarios that correlate with the temporal dynamics of an insect outbreak. Patches represent areas defoliated to the three levels of intensity indicated in panel E. *X*'s indicate hypothetical sample locations. In panel A, the outbreak is at or near its peak. Due to the high population density and density-dependent migration, we expect panmixia at the scale of the study area and a low level of genetic structure (e.g. F_{st}). Panel B shows an intermediate level of spatial genetic structure where populations are not fully connected as one expects during the increasing (towards A) or declining (towards C) phase of an outbreak. Sublabels 1 and 2 indicate two processes involved in the temporal dynamics of spatial genetic structure that depend on whether the outbreak is decreasing or increasing: (i) dispersal among neighbouring populations and the homogenization of genetic structure through admixture; (ii) patch contraction and the creation of genetic structure through isolation of patches. Panel C shows a low population density situation wherein patches are poorly connected and have contracted significantly resulting in well differentiated subpopulations and a high degree of genetic structure. Depending on when these locations would be sampled (pre- or post-peak), processes of patch contraction, expansion or dispersal may be operating. Finally, in panel D, we show that these different scenarios may all be present within a single study region at a larger spatial scale (note difference in scale bar).



synchronization can manifest itself as 'epicentres' spreading to adjacent areas (Berryman *et al.* 1987; Wallner 1987; Liehold & McManus 1991) or as 'travelling waves' where outbreaks spread across large geographic areas (Bjørnstad *et al.* 2002; Tenow *et al.* 2013). The relative importance of dispersal to the synchronization of insect population cycles is equivocal (Royama *et al.* 2005) and will remain so until it can be quantified more precisely (Myers & Cory 2013).

Increasingly, indirect genetics-based methods are used to characterize dispersal (Broquet & Petit 2009; Segelbacher *et al.* 2010). Genetic approaches can be challenging for cyclic populations because dispersal patterns can be confounded by demographic processes. Genetic structure, from which gene flow and dispersal are inferred, is expected to repeatedly increase and diminish as irruptions grow and collapse (Wright 1931; Nei *et al.* 1975), making it difficult to describe patterns that arise exclusively due to dispersal. We hypothesize that the population contractions exhibited in cyclic irruptive systems will be identifiable in spatial genetic structure at regional scales due to the

influence of founder effects as remnant populations enter a period of endemic (in contrast to epidemic) dynamics (Box 1, Franklin *et al.* 2014). This dynamic pattern of genetic differentiation depends on interactions between cycle amplitude and duration as well as context-dependent dispersal (Ehrich *et al.* 2001; Franklin *et al.* 2014).

To test our hypotheses regarding the influence of dispersal and population cycles on spatial genetic structure, we examined spatial patterns of genetic variation of the spruce budworm in a large (69 000 km^2) landscape between Ontario, Canada and Minnesota, U.S.A., hereafter referred to as Border Lakes landscape (BLL; Fig. 1). Our first objective was to test the hypothesis that the legacy of the previous outbreak created spatial genetic structure in the spruce budworm at the scale of the BLL. Given spatial genetic structure, we then set out to test the factors responsible for that structure and in particular whether it could be attributed to immigrants (moths), residents (larvae) or both. If immigrants were implicated, we also set out to determine their geographic source. In comparing resident to migrant spatial genetic structure, we examine the temporal dynamics of spatial genetic structure over effectively two generations. Answers to these questions will provide insights into how dispersal, local and regional demography, and outbreak history interact to shape observable spatial genetic structure and spatial synchrony in cyclic systems. Our results are coupled with insights from more traditional phenological inference of dispersal, as well as statistical analyses and simulation modelling of long-distance movement (Anderson & Sturtevant 2011; Sturtevant *et al.* 2013).

Methods

Study area

The BLL is located at the transition zone between the Great Lakes-St. Lawrence mixed-wood and boreal forest regions and contains large areas with divergent forest disturbance histories created by political boundaries (Sturtevant *et al.* 2014). The central portion of this landscape is a large (1×10^6 ha) unmanaged wilderness area comprised of the Boundary Waters Canoe Area, Voyageurs National Park, and Quetico Provincial Park. Surrounding this wilderness are managed landscapes in the United States and Canada (Anderson & Sturtevant 2011; James *et al.* 2011b; Robert *et al.* 2012). The last widespread outbreak (1985–1991) affected all but the most southwest region of the BLL. A localized outbreak has been continually active at a low intensity in this area since 2003 (Fig. 1).

Insect collection

Spruce budworm larvae (4th–6th instar) were collected over a 2-week period in June 2007 and grouped to seven locations (Fig. 1, Table 1). Specimens were preserved in ethanol and stored at −20 °C. Male adult moths were collected from the same sites using three pheromone traps per location (see Anderson & Sturtevant 2011 for details). Hereafter, we use the term 'adults' to describe male moths. Moths were collected from traps every 7–10 days during the main flight period (late June through mid-July) and a final collection at the end of August. Larvae and moths that were collected are of the same generation. However, the genetic structure of larvae represents the consequences of the previous summer's recruitment that includes both locally produced individuals and immigrants, while the genetic structure of moths sampled at any given location can also include immigrants from the current year.

Genetic data

In total, 237 individuals were collected from 20 locations that were aggregated into seven populations (Fig. 1). Twenty-eight microsatellite loci were assessed

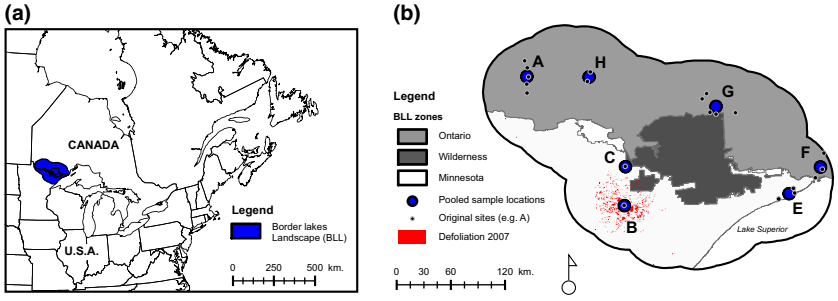


Fig. 1 Study area and context. (a) Continental context of the Border Lakes landscape (BLL). (b) Sample locations and separate regions within the BLL as well as defoliation in 2007. Large blue circles represent centroids of the minimum convex polygon surrounding the original sample locations (small circles). Defoliation is based on aerial surveys conducted in 2007 and was obtained from the U.S. Forest Service (http://na.fs.fed.us/fhp/ta/av/index.shtm; accessed March 2008.)

Table 1 Summary of sample sites, number of individual spruce budworm collected at each site at each sampling period and genetic diversity measures (heterozygosity) for grouping based on larvae vs. adults (grouping = 'life stage') as well as adults (moths) captured at different phenological timings (grouping = 'phenology'). Sixteen loci were typed for each sample location. Values in parentheses indicate standard deviations

Site	Long.	Lat.	Life -stage	Phenology of adult capture	Label 1*	Label 2†	*n*	Unbiased Hz‡	Obs Hz‡	No. alleles
A	−93.929	49.056	Larvae	—	A_L	A_L	5	0.50 (0.08)	0.41 (0.06)	3.19 (1.91)
			Adult	All	—	A_A	9	0.55 (0.06)	0.46 (0.04)	4.31 (3.09)
			Adult	Early	A_Ae	—	3	0.50 (0.10)	0.50 (0.07)	2.75 (1.69)
			Adult	Mid	A_Am	—	6	0.53 (0.06)	0.44 (0.05)	3.63 (2.00)
B	−92.464	47.773	Larvae	—	B_L	B_L	17	0.51 (0.07)	0.40 (0.03)	5.13 (3.91)
			Adult	Mid	B_Am	B_A	17	0.53 (0.06)	0.47 (0.03)	5.31 (4.59)
C	−92.443	48.158	Larvae	—	C_L	C_L	15	0.50 (0.07)	0.38 (0.03)	4.75 (3.75)
			Adult	Mid	C_Am	C_A	16	0.52 (0.07)	0.49 (0.03)	5.13 (3.86)
E	−90.025	47.852	Larvae	—	E_L	E_L	11	0.55 (0.06)	0.40 (0.04)	4.31 (2.68)
			Adult	All	—	E_A	30	0.51 (0.07)	0.47 (0.02)	6.31 (5.52)
			Adult	Early	E_Ae	—	13	0.52 (0.07)	0.48 (0.03)	5.19 (4.15)
			Adult	Mid	E_Am	—	17	0.50 (0.07)	0.46 (0.03)	5.00 (4.13)
F	−89.541	48.105	Larvae	—	F_L	F_L	7	0.53 (0.07)	0.40 (0.05)	3.94 (2.41)
			Adult	All	—	F_A	32	0.52 (0.07)	0.49 (0.02)	6.75 (4.97)
			Adult	Early	F_Ae	—	16	0.52 (0.07)	0.49 (0.03)	5.50 (3.86)
			Adult	Mid	F_Am	—	16	0.52 (0.07)	0.48 (0.03)	5.69 (4.36)
G	−91.073	48.746	Larvae	—	G_L	G_L	11	0.51 (0.07)	0.40 (0.04)	4.25 (2.91)
			Adult	All	—	G_A	30	0.51 (0.07)	0.46 (0.02)	6.75 (6.10)
			Adult	Early	G_Ae	—	14	0.52 (0.07)	0.48 (0.03)	5.13 (4.16)
			Adult	Mid	G_Am	—	16	0.50 (0.07)	0.45 (0.03)	5.44 (4.65)
H	−92.989	49.055	Larvae	—	H_L	H_L	7	0.48 (0.07)	0.41 (0.05)	3.25 (1.69)
			Adult	All	—	H_A	24	0.49 (0.07)	0.46 (0.03)	5.44 (4.99)
			Adult	Early	H_Ae	—	8	0.49 (0.07)	0.45 (0.04)	3.75 (2.77)
			Adult	Mid	H_Am	—	16	0.50 (0.07)	0.46 (0.03)	4.94 (4.23)

*Labels associated with Fig. 3a; grouping = 'phenology'.
†Labels associated with Fig. 3b; grouping = 'life stage'.
‡Nei's Unbiased Heterozygosity (Nei 1973).

for their suitability to describe genetic structure. Two sets of loci were included in these original 28: (i) a set of anonymous genomic loci (Lumley *et al.* 2009) and (ii) a set derived from expressed sequence tag libraries (Brunet *et al.* 2013). DNA was extracted using QIAGEN DNEasy Blood and Tissue extraction kits and PCR-amplification and fragment analysis of each locus followed Brunet *et al.* (2013).

All 28 loci were not suitable for analysis. Loci were removed from the data set using the following criteria: (i) loci that had many (>5) nontyped individuals; (ii) loci that were fixed, (mean dominant allele frequency >0.95) across all populations; and (iii) loci that had an average estimated null allele frequency >0.1 over all populations. Null allele frequencies were estimated using ML-Null (Kalinowski & Taper 2006). Individuals with multiple (>5) nontyped loci (i.e. zero-values) were also removed. The final data set included 231 individuals and 16 loci. Individuals were genotyped once, unless the original run failed in which case a second attempt was made. Automated binning was used in GENEMAPPER 4.0 (ABI) for initial genotyping. All stutter bands for all loci and individuals were assessed by eye to verify that there were no errors in allele calls. Deviation from HWE was assessed using the *hwe.test* function in the ADEGENET package in R (Jombart 2008). Linkage disequilibrium between all pairs of loci was tested using FSTAT 2.9.3.2 (Goudet 2001). Standard diversity measures were calculated for each site using the Excel Microsatellite Toolkit (Park 2001). Raw genotypic data were converted to allele frequencies and geo-referenced to the centroid of the minimum convex polygon surrounding the sites that comprise each sample location (Fig. 1, Table 1). Pairwise F_{st} (Nei 1973) values were calculated using these frequencies for subsequent analysis using the F_{st} function in ADEGENET.

Genetic data were summarized in three ways. First, we summarized larval and adult individuals on the basis of the site at which they were captured without regard for life stage or whether collection matched moth emergence phenology. Second, we grouped individuals by geographic site and separated larvae from

adults (grouping = 'life stage'). Finally, we separated adults from larvae and then further separated adults on the basis of the timing of their capture relative to local phenology (grouping = 'phenology'). That is, we distinguished between adults captured prior to their expected occurrence ('early' adults), and those that were captured during the expected period of emergence ('peak' adults). Early adults are assumed to be immigrants. Expected phenological timing was calculated using a budworm-specific phenology model (Régnière & You 1991) implemented in BioSIM (Régnière & St-Amant 2008).

Genetic structure

Genetic structure was assessed at different levels of organization using F_{st}. Global F_{st} was calculated for five data sets: (i) all data with no differentiation between life stages or timing; (ii) all data separated by life -stage (grouping = 'life stage'); (iii) all data separated by life -stage that also distinguished timing of adult capture (grouping = 'phenology'); (iv) larvae only; and (v) adults only. Estimates of global F_{st} and 95% confidence intervals were calculated using FSTAT 2.9.3.2 (Goudet 2001). Principal coordinates analysis (PCoA) was used to visualize genetic differentiation in the second and third data sets described above. To determine whether adults and larvae differed, we assessed concordance using a Mantel test and a Procrustes rotation test on their respective F_{st}-based ordination solutions (PCoA) (Peres-Neto & Jackson 2001; James *et al.* 2011a). All ordinations were performed using the *capscale* function in the VEGAN package in R.

Genetic structure was assessed using STRUCTURE v.2.3.4 (Pritchard 2000; Falush *et al.* 2003). This programme uses a Bayesian algorithm to assign individuals to one or more clusters while maintaining HWE and linkage disequilibrium at each locus to identify the optimal number of clusters (K). We assumed admixture and correlated allele frequencies among locations as dispersal and gene flow are expected in this system (Falush *et al.* 2003). For each value of K (1–10), 500 000 Markov chain Monte Carlo (MCMC) generations were sampled following a 50 000 generation burn-in period. MCMC sampling was replicated 10 times for each K. The most likely number of genetic clusters was determined by the 'highest ΔK method' (Evanno *et al.* 2005) and by comparison of mean log probability for each K. Separate structure analyses were conducted for each of five data groupings described above. Data groupings 1, 3 and 4 were also analysed using sampling location as a prior (Hubisz *et al.* 2009). Results of independent replicates were summarized using CLUMPP v.1.1.2b (Jakobsson & Rosenberg 2007).

Association between STRUCTURE clusters and sample sites

We used a chi-square test of association to assess the degree to which the clusters identified by STRUCTURE were correlated with sample locations. More specifically, we tested whether the assignment of individuals to each cluster differed significantly from what one would expect under a null hypothesis of panmixis where individuals from all geographic sites would be found at similar frequencies in the identified clusters. Chi-square analysis was performed using the *chisq.test* function in R.

Isolation by distance

We tested for isolation by distance (IBD) using a Mantel test of correlation between linearized genetic distance ($F_{st}/1-F_{st}$) and log-transformed geographic distance matrices (Rousset 1997). IBD was assessed for all data as well as larvae and adults separately. Significant correlation between larval and moth genetic distances was assessed using a symmetric Procrustes rotation test and a Mantel test using functions in VEGAN in R. A partial Mantel test was used to assess the same relationship while controlling for the effects of geographic distance among sites.

Alternative connectivity hypotheses

In the absence of clear patterns of spatial structure and IBD (see below), we hypothesized that there may be more complicated patterns of genetic similarity among sites. We developed several a priori hypotheses (Tables 2 and 3) to describe genetic connectivity among sites and tested them using constrained ordination. Hypotheses were developed for life -stage ('L'-Hypotheses; Table 2) and phenology ('P'-Hypotheses; Table 3) groupings and were based on: (i) geography (hypotheses 1 and 2); (ii) assumed differences between adults and larvae (hypothesis 3); (iii) known areas of epidemic activity (sites B and C) that were assumed to represent source populations (hypothesis 4); and (iv) an elaboration of hypothesis 4 that includes information from initial ordinations (Fig. 3a) as several sites (e.g. Ae, Am) were more differentiated than expected (hypothesis 5; grouping = 'phenology').

Models were tested using distance-based redundancy analysis (*db*-RDA; Legendre & Anderson 1999) implemented using the *capscale* function in VEGAN in R. Here, the response variable was the double square root-transformed genetic distance matrix (F_{st}). The predictor variables were categorical classifications of each location according to our hypotheses, converted into a model matrix of dummy variables. This approach implements decomposition of variance based on different groupings

Table 2 Summary of connectivity hypotheses examined using groupings that differentiate larvae and moths and include moth phenology ('P'-Hypotheses; grouping = 'phenology'). Hypotheses were tested using *db*-RDA. Actual coding of hypotheses can be found in Table S2 (Supporting Information)

Hypothesis	Description
P1	Coarse regional groupings. The study area was divided into four groups: (i) A and H; (ii) B and C; (iii) E and F; (iv) G. See Fig. 1
P2	Simple/local geographic hypothesis. Each site is considered as its own, regardless of moth or larval forms
P3	Insect life-stage hypothesis. All moths are one group, regardless of phenology, and larvae are another
P4	Epidemic-source hypothesis based on known defoliation at sites B and C. Early moths at all sites are grouped with all samples from B and C. Larvae and peak moths from all other sites are assumed to be independent
P5	Epidemic-source hypothesis (hypothesis 4) modified using initial ordination results (see Fig. 3a). Sites B, C, E and F are grouped together regardless of life stage or phenology. Larvae from sites A, G and H are independent. Early moths are grouped together along with larvae from sites B and C. Mid moths from A and early moths from H are grouped

Table 3 Summary of connectivity hypotheses examined using groupings based on life stage only ('L'-Hypotheses; adults vs. larvae), without regard for phenology. Actual coding of hypotheses can be found in Table S3 (Supporting Information)

Hypothesis	Description
L1	Coarse regional groupings. The study area was divided into four groups: (i) A and H; (ii) B and C; (iii) E and F; (iv) G
L2	Simple geographic hypothesis. Each site is its own, regardless of moth or larval forms
L3	Insect life-stage hypothesis. Adults are one group, larvae another
L4	Epidemic-source hypothesis based on known defoliation activity at sites B and C. Moths and larvae from sites B and C are grouped together. All other larvae are independent. All other moths are grouped with sites B and C

similar to AMOVA (Excoffier *et al.* 1992) but using constrained ordination. The genetic response matrix was transformed to ensure that it was Euclidean for use in *db*-RDA. Specific coding of all hypotheses is shown in Supporting Information.

Results

Genetic diversity

Little variation was observed in either heterozygosity or number of alleles among sites. Genetic diversity measures for both groupings ('life form' and 'phenology') are shown in Table 1. Multiple loci were not in HWE according using a Bonferroni-corrected significance threshold at the two different levels of organization. Of the 16 loci used, the proportion that was not in HWE was uniformly distributed among sample locations (Tables S5 and S6, Supporting Information). No significant linkage disequilibrium was detected.

Genetic structure

Overall, significant genetic structure was not detected. Global F_{st} for all samples when pooling adults and larvae was weak and nonsignificant (F_{st} = 0.002, SE = 0.002, n = 7). Global F_{st} for the data arranged such that larvae and adults represented unique sites (albeit sharing the same geographic location) was 0.006 (SE = 0.003, n = 14). When adults were further separated based on the phenology of their capture and distinct from larvae, genetic structure remained weak (F_{st} = 0.005, SE = 0.003, n = 19). Finally, genetic structure was low and nonsignificant for adults (F_{st} = 0.006, SE = 0.003, n = 7) and larvae analysed independently (F_{st} = 0.005, SE = 0.003, n = 7).

Population genetic structure was not detected using STRUCTURE for any of the data groupings that included adults (not shown). However, when examining larvae alone, both ΔK and the mean log probability of K gave support for the existence of some structure. ΔK values supported two (4.03) and five (3.47) genetic clusters; the highest mean log probability was for K = 5 (−2687.07). The use of location priors gave further support for K = 5 clusters (ΔK = 17.03; mean log probability = −2646.91; Fig. 2). Partitioning of individuals at K = 2 with the straight admixture model found no clustering of individuals from similar locations; all individuals have nearly equal assignment probability to each of the two clusters. Differences between analyses with

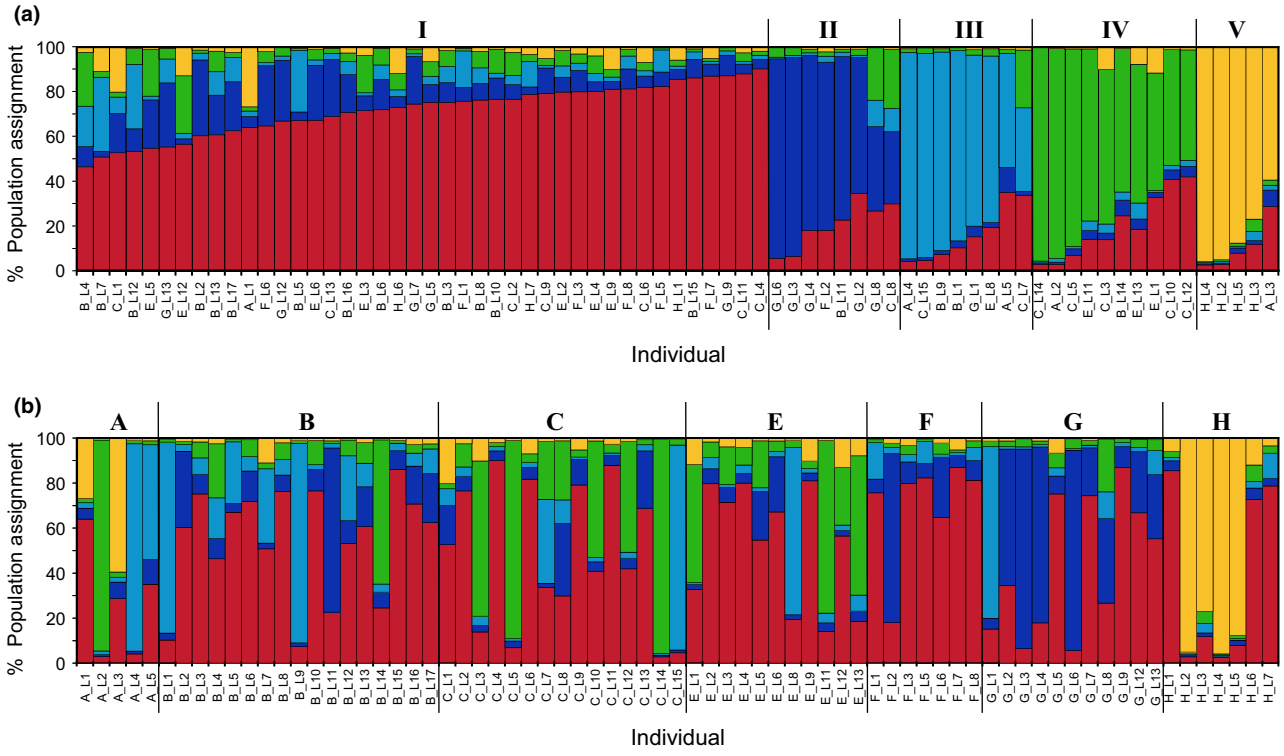


Fig. 2 STRUCTURE results using geographic location as a prior (K = 5). Colours in each column indicate the probability of membership of an individual larva to each of the five clusters. (a) shows larvae organized according to cluster membership (I–V), while (b) shows larvae organized according to geographic location (A–H). Clusters II, III, IV and V were, respectively, associated with sites G, A, C and H more than expected, whereas cluster I was associated with all sites (see Table 4).

Table 4 χ^2 table to test the assignment of predefined geographic populations (A–F) to clusters identified using STRUCTURE (I–V). Each cell represents an individual χ^2 value. Most important individual contributions to the overall χ^2 statistic are in bold

	Clusters				
	I	II	III	IV	V
Sites					
A	1.22	0.55	**3.85**	0.14	1.26
B	1.06	0.40	0.01	0.76	1.16
C	0.31	0.25	0.08	**4.22**	1.03
E	0.07	1.21	0.04	1.48	0.75
F	0.97	0.07	0.77	0.96	0.48
G	0.28	**11.94**	0.04	1.51	0.75
H	0.26	0.77	0.77	0.96	**25.85**

Total χ^2: 66.21; d.f. = 24; $P < 0.005$

and without priors were only evident in the number of individuals assigned to each cluster; most individuals had high levels of admixture among the identified five genetic clusters.

Although we did not find strict correspondence between sample sites and genetic clusters, some interesting patterns emerged. Overall, individual membership in the five clusters identified by STRUCTURE was not randomly distributed across all geographic populations (χ^2 = 66.21; d.f. = 24; P = 0.0005). In examining individual contributions of the χ^2 values (Table 4), we notice that the northern sites G, A, C and H are more strongly associated with clusters II, III, IV and V, respectively, and that the southern sites B, E and F are not uniquely associated with any cluster. Clusters II, III, IV and V may represent local genetic structure that was created following the collapse of the previous outbreak (see Box 1).

Multivariate analysis

Principal coordinates analysis of the F_{st} distance matrix showed interesting variation among sites that remained present when we included the phenological timing of adult capture (Fig. 3). When considering larvae and adults separated by phenology (Fig. 3a), it appears that some sites are more similar than others and that larvae appear more differentiated, in general, than adults. The first and second ordination axes captured 13.3% and 9.8% of the variation in allele frequencies, respectively. When considering adults and larvae separately without regard for adult phenology (Fig. 3b), adults appear

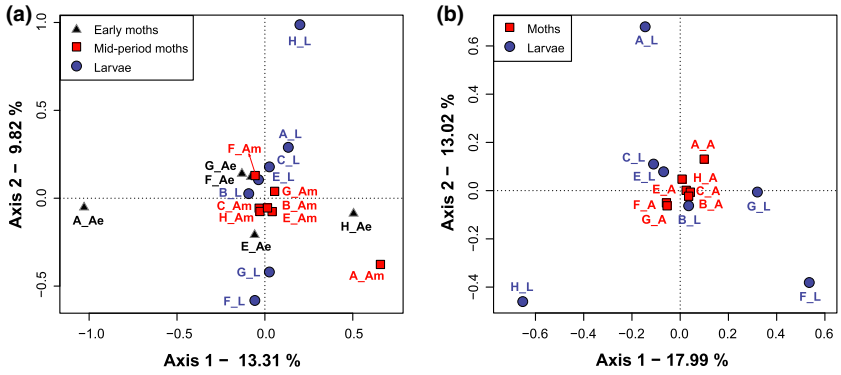


Fig. 3 PCoA ordinations of genetic distances (F_{st}) among sample sites. (a) Larvae and moths are separated, and moths are further divided based on phenology of moth collection (grouping = 'phenology'). Early moths refer to those collected before their predicted phenology, whereas mid-period moths were those collected when the phenological model (BioSIM) predicted moth presence. (b) shows separation of samples based only on adults vs. larvae (grouping = 'life stage'). Site codes in both panels correspond with those described in Table 1.

more similar to each other than to larvae and indicate a common source population from which the adults likely originated (e.g. site B). In this case, the first and second ordination axes captured 18% and 13% of the variation in allele frequencies.

Isolation by distance

Significant IBD was not detected for larval, adult or combined data groupings (Fig. 4; Table S1, Supporting Information). Larval and adult genetic distance matrices were also not correlated, even when controlling for the effect of geographic distance (Mantel test; Table S1, Supporting Information). This nonsignificant correlation was corroborated by a Procrustes rotation test of the F_{st}-based PCoA ordination solutions for larvae and adults ($r = 0.65$; $P = 0.149$). Genetic distances (F_{st}) between a few larval sites were greater than the rest including comparisons between sites A, H, G and F, but not sites B and C. This corresponds well with our epidemic-source hypotheses described below for which we found support (Tables 5 and 6). Removal of these more distinguished comparisons did not change the overall result of nonsignificant slopes.

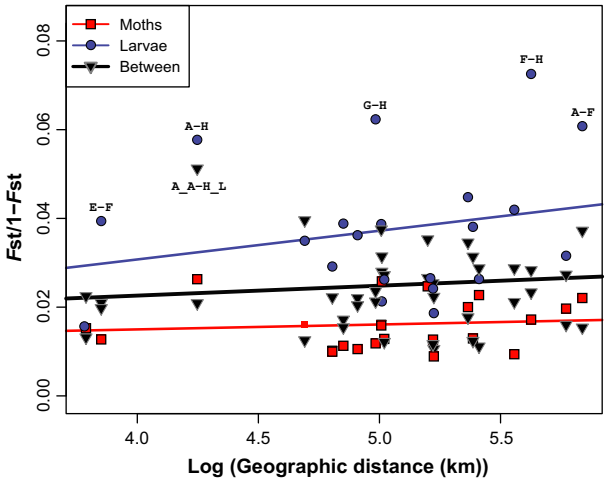


Fig. 4 Isolation-by-distance (IBD) plot of genetic distance (F_{st}) by geographic distance for all sites including comparison within larval (blue circles) and moth (red squares) groupings. Black triangles represent pairwise comparisons between life stages (i.e. moth vs. larvae). Trend lines represent linear regressions between geographic and genetic distances for the separate life stages (blue and red lines) as well as the relationship between all samples (black thick line). Significance of these relationships was tested using Mantel tests; none was significant (see Table S1, Supporting Information). Sites with higher than average F_{st} values are labelled for interpretation; see text for details.

Alternative connectivity hypotheses

In testing both the 'P' and 'L' hypotheses, we found support for only the epidemic-source model presumably driven by a patch of ongoing defoliation at sites B and C (Fig. 1; Tables 5 and 6). For groupings that accounted for phenology, the simple epidemic-source hypothesis was not supported (hypothesis P4) but a slightly elaborated model (hypothesis P5) informed by initial ordinations was (Fig. 3a; Table 5). For groupings that did not account for phenology, we found the strongest support for the model in which all adults were grouped with larvae in sites B and C (hypothesis L4); all other larvae were assumed to be independent. These models also returned the highest adjusted-R^2 (0.16 and 0.17 for hypotheses P5 and L4, respectively) and lowest Akaike's Information Criterion (AIC), indicating that these hypotheses are more plausible than the others tested (Tables 5 and 6).

Table 5 Results from multiple distance-based redundancy (*db*-RDA) models to describe observed spatial genetic variation between larvae and moths, including moth phenology. Models represent hypotheses described in Table 2 ('P'-Hypotheses). Significant results are indicated in bold

Model	d.f.	F	P	R^2	R^2_{adj}	AIC
P1	3	0.783	1.000	0.135	−0.038	12.321
P2	6	0.887	0.970	0.307	−0.039	14.107
P3	1	1.090	0.170	0.060	0.005	9.903
P4	5	1.090	0.180	0.295	0.024	12.43
P5	**6**	**1.590**	**0.010**	**0.443**	**0.164**	**9.971**

Table 6 Results from multiple distance-based redundancy (*db*-RDA) models to describe observed spatial genetic variation between larvae and moths. Models represent hypotheses described in Table 3 ('L'-Hypotheses). Significant results are indicated in bold

Model	d.f.	F	P	R^2	R^2_{adj}	AIC
L1	3	0.967	0.640	0.225	−0.008	3.135
L2	6	1.028	0.380	0.468	0.013	3.854
L3	1	0.651	0.990	0.051	−0.028	1.962
L4	**5**	**1.546**	**0.005**	**0.491**	**0.174**	**1.234**

Discussion

We presented a snap-shot description of population genetic structure over a large landscape in the spruce budworm, an economically important cyclic species. The determinants of cyclic population dynamics and scale of spatial synchrony in this species have been a focus of basic and applied ecological research for almost a century (Swaine & Craighead 1924). At the scale of our study area, we found that larvae and adults exhibited different spatial patterns in genetic variation and that in general adults appear to originate from a few sites (e.g. site B; Fig. 3). This finding corresponds well with observations of a developing outbreak around site B in 2007 during our collection period which served as a source of migrants for other populations (i.e. the epidemic-source hypothesis). That larvae and adults are not significantly correlated suggests that at the scale of our study area, dispersal plays a very important role in population dynamics and the spatio-temporal apportionment of genetic variation. Using STRUCTURE, we found that larvae could be separated into five clusters and that four of these clusters correspond in part to geographic locations (Fig. 2, Table 4). The fifth larval cluster was comprised of individuals found across all sites. In contrast, adults were consistently assigned to a single group (not shown). Because structure was identified among larvae and not adults, this suggests that structure is likely spatially and temporally dynamic. Through the combined forces of dispersal, drift and demographic variation, we would expect this structure to become either strengthened or diminished depending on the ecological context of outbreak progression and population dynamics as described in Box 1.

Although Harvey (1996) concluded that the spruce budworm was panmictic over a very large portion of its native range, we identified interesting spatial patterns in larval genetic structure at a smaller spatial scale. The identified patterns were contingent on our separation of life -stages which allowed us to contrast the spatial genetic patterns in residents (larvae) and migrants (adult males). The generally higher variability in microsatellite loci relative to isozymes used previously (Harvey 1996) may have increased our ability to detect contemporary genetic structure. However, had we sampled 5 years previously, or 5 years subsequently, it is possible we would have found different patterns of genetic structure. Our conclusions are necessarily restricted to this particular ecological context.

Inference from population genetic analysis depends on awareness of the multiple ecological, demographic and micro-evolutionary processes that create and diminish spatial genetic structure through time (Box 1). Large outbreaks in the centre of the budworm outbreak range go through all the stages shown in Box 1. In contrast, regions on the geographic periphery of this range (e.g. the southern portion of the BLL) are characterized by smaller scale local outbreaks. These outbreaks tend not to coalesce, nor do they reach the 'red' stage (Box 1, Panel A). Adjacent and typically larger outbreaks to the north become a significant source of emigrants. Smaller outbreaks (e.g. Box 1, Panel B) can also function as a source, but to a lesser degree, which can still confound our ability to assess the 'scale of spatial genetic structure' and effective dispersal. The localized outbreak in the BLL is an example of such an intermediate outbreak in which connectivity and dispersal can still influence genetic structure but not to the same scale, intensity and duration one would expect in more continuous central regions of the spruce budworm range.

STRUCTURE results (Fig. 2) differed from the PCoA (Fig. 3) because individuals were grouped in the ordination analysis, whereas STRUCTURE is based on individuals. Using an individual-based analysis, we found that most geographic sites contain a mix of individuals from different genetic clusters. Follow-up χ^2 analysis (Table 4) indicated that a single cluster was present at all sites (cluster I) and that four other clusters (II, III, IV and V) showed moderate fidelity to four unique sites. The three sites that were not specifically associated with any cluster included site B at the centre of the outbreak as well as sites E and F that are the most closely aligned

with prevailing downwind direction from site B (Anderson & Sturtevant 2011). It is tempting to speculate that cluster I may represent an emerging outbreak 'type' that is found at all sites and originates from site B. Clusters II–V could then represent the legacies of the previous outbreak's collapse and contraction (Box 1) whose unique genetic profile will likely be soon overcome by immigrants from the source population at site B, much like sites E and F. However, rigorous testing for such a hypothesis will require additional sampling and more precise molecular markers.

Genetic methods have been applied in several other cyclic systems to characterize dispersal. In the canonical cyclic system of the Canadian Lynx (*Lynx canadensis*) (Moran 1953), dispersal has been found to be more important than drift in determining genetic structure (Schwartz *et al.* 2002; Row *et al.* 2012). Similarly, in comparing populations of two lemming species, differences in dispersal were found to account for differences in the spatial scale of genetic structure (Ehrich *et al.* 2001). The effects of demographic bottlenecks and spatial connectivity on the population genetics of noncycling Lepidoptera have also been previously examined (Hastings & Harrison 1994; Ehrlich & Hanski 2004; Keyghobadi *et al.* 2005). Recently, these conceptual approaches were applied to the cycling western tent caterpillar where high levels of dispersal and fluctuating dynamics at a relatively fine temporal scale accounted for the lack of genetic differentiation among populations (Franklin *et al.* 2014). Other studies of cyclic Lepidoptera have found results that differ with the molecular marker used. In the winter moth (*Operophtera brumata*), genetic structure was identified using allozymes in Belgium (Van Dongen *et al.* 1998), whereas no structure was found using AFLPs in the UK (Leggett *et al.* 2011). In both the western tent caterpillar and winter moth system, populations cycle regularly across 3–4 orders of magnitude every 9–11 years, which may be too short a time period to permit spatial genetic differentiation (Norén & Angerbjörn 2014). In contrast, we examined 16 microsatellite loci in the spruce budworm, whose populations can vary over 4–6 orders of magnitude in abundance every *c.* 35 years in the central part of its range (Jardon *et al.* 2003). Although spruce budworm moths display a high propensity for migration, the long period of low population density between outbreaks reduces effective dispersal (Régnière & Nealis 2007) and may permit sufficient genetic differentiation, even though our study area is found at the southern edge of this range. The duration of this endemic period likely affects the rates of genetic differentiation and our ability to identify structure (Box 1). Our capacity to identify genetic structure in this context may also depend on the marker used (Ellegren 2004) as suggested by the continental-scale panmictic conclusion of previous work on the spruce budworm using isozymes (Harvey 1996). Future work will focus on improving the characterization spruce budworm genetic variation using next-generation sequencing and SNP markers developed using a genotyping-by-sequencing approach (Elshire *et al.* 2011).

Conclusions

Characterizing genetic structure to infer dispersal is challenging (Broquet & Petit 2009). Cyclic populations present additional complications that must be considered including: (i) cycle periodicity and amplitude; (ii) dispersal capacity; and (iii) ecological context (i.e. local and regional outbreak status; Box 1). In highly mobile taxa, such as irruptive Lepidoptera, life stage (e.g. mobile vs. sessile) and phenological timing of sample collection relative to local outbreak status may also be important. We identified patterns of genetic variation in the spruce budworm that indicate a landscape in flux in which current genetic structure is potentially transient rather than stable. Interpretation of these patterns relied upon specific knowledge of current regional outbreak status, local populations irruptions and regional wind patterns that determine which sites were receiving immigrants (Anderson & Sturtevant 2011; Sturtevant *et al.* 2013). In examining a system that includes a known localized outbreak population and using information on the phenology of moth emergence, we were able to distinguish residents from migrants and gained insight into the processes underlying population connectivity and spatial synchrony in this system.

Ecologists have long considered the effects of demographic bottlenecks on the spatial genetic structure of cyclic populations (Elton 1924). The dynamics and resultant genetic structure in such systems can help us to understand the role of the multiple processes that lead to spatial synchrony in population irruptions, including dispersal. Dispersal is an essential component of spatial population dynamics that is modulated by spatial heterogeneity (Kareiva *et al.* 1990). Indeed, it is the interactions between dispersal and spatial heterogeneity that encapsulate the essential question of how and why populations cycle (Liebhold *et al.* 2004; Myers & Cory 2013). Further investigations into population dynamics using spatial and temporal genetic analysis can help us to better understand the mechanisms that affect both the amplitude and frequency of population cycles at regional and local spatial scales.

Acknowledgements

This research was supported by the USDA Cooperative State Research, Education, and Extension Service (CSREES) Managed

Ecosystems program (2005-35101-16342) and the US National Fire Plan. FS, MJF, and PJ were each supported by a National Science and Engineering Research Council Discovery Grant (Canada). We thank Dean Anderson for his role in the collection of field samples. Leah Berkman, Michel Cusson and two anonymous reviewers provided helpful comments on prior versions of the manuscript.

References

Anderson DP, Sturtevant BR (2011) Pattern analysis of eastern spruce budworm, *Choristoneura fumiferana*, dispersal. *Ecography*, **34**, 488–497.

Berryman AA (1987) The theory and classification of outbreaks. In: *Insect Outbreaks* (eds Barbosa P, Schultz JC), pp. 3–30. Academic Press, San Diego.

Bjørnstad ON, Bascompte J (2001) Synchrony and second-order spatial correlation in host–parasitoid systems. *Journal of Animal Ecology*, **70**, 924–933.

Bjørnstad ON, Ims RA, Lambin X (1999) Spatial population dynamics: analyzing patterns and processes of population synchrony. *Trends in Ecology & Evolution*, **14**, 427–432.

Bjørnstad ON, Peltonen M, Liebhold AM, Baltensweiler W (2002) Waves of larch budmoth outbreaks in the European Alps. *Science*, **298**, 1020–1023.

Broquet T, Petit EJ (2009) Molecular estimation of dispersal for ecology and population genetics. *Annual Review of Ecology, Evolution and Systematics*, **40**, 193–216.

Brunet B, Doucet D, Sturtevant B, Sperling F (2013) Characterization of EST-based SSR loci in the spruce budworm, *Choristoneura fumiferana* (Lepidoptera: Tortricidae). *Conservation Genetics Resources*, **5**, 541–544.

Ehrich D, Krebs CJ, Kenney AJ, Stenseth NC (2001) Comparing the genetic population structure of two species of arctic lemmings: more local differentiation in *Lemmus trimucronatus* than in *Dicrostonyx groenlandicus*. *Oikos*, **94**, 143–150.

Ehrlich PR, Hanski I (Eds) (2004) *On the Wings of Checkerspots: A Model System for Population Biology*. Oxford University Press, Oxford, UK.

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics*, **5**, 435–445.

Elshire RJ, Glaubitz JC, Sun Q *et al.* (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, **6**, e19379.

Elton CS (1924) Periodic fluctuations in the numbers of animals: their causes and effects. *British Journal of Experimental Biology*, **2**, 119–163.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Franklin MT, Myers JH, Cory JS (2014) Genetic similarity of island populations of tent caterpillars during successive outbreaks. *PLoS ONE*, **9**, e96679.

Goudet J (2001) FSTAT, a program to estimate and test gene diversity and fixation indices (Version 2.9.3.2). http://www2.unil.ch/popgen/softwares/fstat.htm.

Greenbank DO, Schaefer GW, Rainey RC (1980) Spruce budworm (Lepidoptera: Tortricidae) moth flight and dispersal: new understanding from canopy observations, radar, and aircraft. *Memoirs of the Entomological Society of Canada*, **110**, 1–49.

Hardy Y, Lafond A, Hamel L (1983) The epidemiology of the current spruce budworm outbreak in Quebec. *Forest Science*, **29**, 715–725.

Harvey G (1996) Population genetics of the spruce budworm, *Choristoneura fumiferana* (Clem.) Freeman (Lepidoptera: Tortricidae), in relation to geographical and population density differences. *The Canadian Entomologist*, **128**, 219–243.

Hastings A, Harrison S (1994) Metapopulation dynamics and genetics. *Annual Review of Ecology and Systematics*, **25**, 167–188.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

James PMA, Coltman DW, Murray BW, Hamelin RC, Sperling FAH (2011a) Spatial genetic structure of a symbiotic beetle-fungal system: toward multi-taxa integrated landscape genetics. *PLoS ONE*, **6**, e25359.

James PMA, Sturtevant BR, Townsend P, Wolter P, Fortin M-J (2011b) Two-dimensional wavelet analysis of spruce budworm host basal area in the Border Lakes landscape. *Ecological Applications*, **21**, 2197–2209.

Jardon Y, Morin H, Dutilleul P (2003) Periodicité et synchronisme des épidemies de la tordeuse des bourgeons de l'epinette au Québec. *Canadian Journal of Forest Research*, **33**, 1947–1961.

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403–1405.

Kalinowski ST, Taper ML (2006) Maximum likelihood estimation of the frequency of null alleles at microsatellite loci. *Conservation Genetics*, **7**, 991–995.

Kareiva P, Mullen AR, Southwood A (1990) Population dynamics in spatially complex environments: theory and data. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, **330**, 175–190.

Keyghobadi N, Roland J, Strobeck C (2005) Genetic differentiation and gene flow among populations of the alpine butterfly, *Parnassius smintheus*, vary with landscape connectivity. *Molecular Ecology*, **14**, 1897–1909.

Krebs CJ, Kenney AJ, Gilbert S *et al.* (2002) Synchrony in lemming and vole populations in the Canadian Arctic. *Canadian Journal of Zoology*, **80**, 1323–1333.

Legendre P, Anderson MJ (1999) Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecological Monographs*, **69**, 1–24.

Leggett HC, Jones EO, Burke T *et al.* (2011) Population genetic structure of the winter moth, *Operophtera brumata* Linnaeus, in the Orkney Isles suggests long-distance dispersal. *Ecological Entomology*, **36**, 318–325.

Liebhold A, Koenig WD, Bjørnstad ON (2004) Spatial synchrony in population dynamics. *Annual Review of Ecology, Evolution and Systematics*, **35**, 467–490.

Liebhold AM, Kyle JH, Ottar NB (2012) Spatial synchrony of insect outbreaks. In: *Insect Outbreaks Revisited* (eds Barbosa P, Letourneau DK, Agrawal AA), pp. 113–125, 1st edn. Wiley-Blackwell, Hoboken, New Jersey.

Liehold AM, McManus ML (1991) Does larval dispersal cause the expansion of gypsy moth outbreaks? *Northern Journal of Applied Forestry*, **8**, 95–98.

Lumley L, Davis C, Sperling F (2009) Isolation and characterization of eight microsatellite loci in the spruce budworm species *Choristoneura fumiferana* and *Choristoneura occidentalis*, and cross-species amplification in related tortricid moths. *Conservation Genetics Resources*, **1**, 501–504.

Moran P (1953) The statistical analysis of the Canadian Lynx cycle. *Australian Journal of Zoology*, **1**, 291–298.

Myers JH, Cory JS (2013) Population cycles in forest Lepidoptera revisited. *Annual Review of Ecology, Evolution, and Systematics*, **44**, 565–592.

Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the USA*, **70**, 3321–3323.

Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution*, **29**, 1–10.

Norén K, Angerbjörn A (2014) Genetic perspectives on northern population cycles: bridging the gap between theory and empirical studies. *Biological Reviews*, **89**, 493–510.

Park S (2001) *The Excel Microsatellite Toolkit (version 3.1)*, Animal Genomics Laboratory, UCD, Ireland.

Peltonen M, Liebhold AM, Bjørnstad ON, Williams DW (2002) Spatial synchrony in forest insect outbreaks: roles of regional stochasticity and dispersal. *Ecology*, **83**, 3120–3129.

Peres-Neto PR, Jackson DA (2001) How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. *Oecologia*, **129**, 169–178.

Pritchard JK (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Régnière J, Nealis V (2007) Ecological mechanisms of population change during outbreaks of the spruce budworm. *Ecological Entomology*, **32**, 461–477.

Régnière J, St-Amant R (2008) *BioSIM 9 User's Manual*. Information Report LAU-X-134. Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre, Quebec City, Quebec, Canada.

Régnière J, You M (1991) A simulation model of spruce budworm (Lepidoptera: Tortricidae) feeding on balsam fir and white spruce. *Ecological Modelling*, **54**, 277–297.

Régnière J, Delisle J, Pureswaran DS, Trudel R (2013) Mate-finding allee effect in spruce budworm population dynamics. *Entomologia Experimentalis et Applicata*, **146**, 112–122.

Robert L-E, Kneeshaw D, Sturtevant BR (2012) Effects of forest management legacies on spruce budworm (*Choristoneura fumiferana*) outbreaks. *Canadian Journal of Forest Research*, **42**, 463–475.

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

Row J, Gomez C, Koen E *et al.* (2012) Dispersal promotes high gene flow among Canada lynx populations across mainland North America. *Conservation Genetics*, **13**, 1259–1268.

Royama T (1984) Population dynamics of the spruce budworm *Choristoneura fumiferana*. *Ecological Monographs*, **54**, 429–462.

Royama T, MacKinnon WE, Kettela EG, Carter NE, Hartling LK (2005) Analysis of spruce budworm outbreak cycles in New Brunswick, Canada, since 1952. *Ecology*, **86**, 1212–1224.

Schwartz MK, Mills LS, McKelvey KS, Ruggiero LF, Allendorf FW (2002) DNA reveals high dispersal synchronizing the population dynamics of Canada lynx. *Nature*, **415**, 520–522.

Segelbacher G, Cushman SA, Epperson BK *et al.* (2010) Applications of landscape genetics in conservation biology: concepts and challenges. *Conservation Genetics*, **11**, 375–385.

Sturtevant BR, Achtemeier GL, Charney JJ *et al.* (2013) Long-distance dispersal of spruce budworm *Choristoneura fumiferana* Clemens in Minnesota (USA) and Ontario (Canada) via the atmospheric pathway. *Agricultural and Forest Meteorology*, **168**, 186–200.

Sturtevant BR, Miranda BR, Wolter PT *et al.* (2014) Forest recovery patterns in response to divergent disturbance regimes in the Border Lakes region of Minnesota (USA) and Ontario (Canada). *Forest Ecology and Management*, **313**, 199–211.

Swaine JM, Craighead FC (1924) Studies on the spruce budworm (*Cacoecia fumiferana* Clem.) Part 1: a general account of the outbreaks, injury and associated insects. Dominion of Canada, Dept of Agriculture, Technical Bulletin No. 37. Ottawa, ON, pp. 3–27.

Tenow O, Nilssen AC, Bylund H *et al.* (2013) Geometrid outbreak waves travel across Europe. *Journal of Animal Ecology*, **82**, 84–95.

Tobin PC, Bjørnstad ON (2003) Spatial dynamics and cross-correlation in a transient predator–prey system. *Journal of Animal Ecology*, **72**, 460–467.

Tobin PC, Bjørnstad ON (2005) Roles of dispersal, stochasticity, and nonlinear dynamics in the spatial structuring of seasonal natural enemy–victim populations. *Population Ecology*, **47**, 221–227.

Van Dongen S, Backeljau T, Matthysen E, Dhondt A (1998) Genetic population structure of the winter moth (*Operophtera brumata* L.) (Lepidoptera, Geometridae) in a fragmented landscape. *Heredity*, **80**, 92–100.

Wallner W (1987) Factors affecting insect population dynamics: differences between outbreak and non-outbreak species. *Annual Review of Entomology*, **32**, 317–340.

Williams DW, Liebhold AM (2000) Spatial synchrony of spruce budworm outbreaks in eastern North America. *Ecology*, **81**, 2753–2766.

Wright S (1931) Evolution in Mendelian populations. *Genetics*, **16**, 97.

P.M.A.J. worked with B.J.C. and B.R.S. to develop hypotheses, led the statistical analysis, and led the writing of the manuscript. B.R.S. led the research team, worked with B.J.C., V.S.Q. and M.J.F. to design the field sampling protocol, led the field effort, designed the study with B.J.C. and wrote sections of the manuscript. B.R.S., P.M.A.J., B.J.C., V.S.C. and M.J.F. collected field specimens. B.J.C. performed the chi-square analysis and wrote sections of the manuscript. V.S.C. and B.R.S.

prepped samples for analysis. B.M.T.B. and L.M.L. performed initial genetic analysis and contributed text from preliminary reports under the direction of F.A.H.S. who coordinated the molecular lab work. B.M.T.B. completed the STRUCTURE analysis. All coauthors contributed to underlying ideas through team discussions and contributed to editing the paper..

Data accessibility

Microsatellite primers are indicated in Brunet *et al.* (2013) and Lumley *et al.* (2009). Genotypes and geographic locations are available on Dryad: doi:10.5061/dryad.jc0df.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Summary of Mantel tests for correlation (*r*) between genetic distance matrices ($F_{st}/1-F_{st}$) for larvae and adult moths, and log-transformed geographic distance.

Table S2 Coding for connectivity hypotheses tested using *db*-RDA including site groupings based on phenology (P-Hypotheses; moths with phenology vs. larvae).

Table S3 Coding for connectivity hypotheses tested using *db*-RDA including site groupings based on life stage only ("L"-Hypotheses; moths vs. larvae).

Table S4 Sources and names of microsatellite loci used. Locus ID refers to labels used in online data accessibility (Dryad).

Table S5 Results from HWE tests for data grouped according to life-stage.

Table S6 Results from HWE tests for data grouped according to phenology.