Genetic Evidence of Broad Spreading of *Lymantria dispar* in the West Siberian Plain

Martemyanov Vyacheslav[1,2]*, Bykov Roman[3], Demenkova Marya[3], Gninenko Yuri[4], Romancev Sergei[5], Bolonin Ivan[5], Mazunin Ilia[6], Belousova Irina[1], Akhanaev Yuri[1], Pavlushin Sergey[1], Krasnoperova Polina[7], Ilinsky Yury[3,6]*

Affiliation

1 Institute of Systematics and Ecology of Animals SB RAS, Novosibirsk. Russia

2 Biological Institute, National Research Tomsk State University, Tomsk, Russia

3 Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

4 All-Russian Research Institute for Silviculture and Mechanization of Forestry, Pushkino, Moscow region, Russia

5 FSI "Russian Centre of Forest Health", Novosibirsk, Russia

6 Baltic Federal University, Kaliningrad, Russia

7 Siberian Federal University, Krasnoyarsk, Russia

* Correspondence authors: V. Martemyanov martemyanov79@yahoo.com;

Yu. Ilinsky paulee@bionet.nsc.ru

Abstract

Gypsy moth *Lymantria dispar* L. 1758 (Lepidoptera: Erebidae) is one of the most dangerous forest pests of the Holarctic region. Outbreaks of gypsy moth populations lead

to significant defoliation of local forests. Within the vast territory of the West Siberian Plain, we noted the movement of the outbreak front in the north-east direction with a speed of approximately 100-200 km per year. The reason for the outbreak's movement is still unclear because *L. dispar* females are characterised by flight ability, which is not enough to provide this movement *per se*. Herein, we analysed the mtDNA pattern of *L. dispar* populations collected from the vast territory of the West Siberian Plain to determine the boundaries of populations and reveal the effect of the outbreak's front movement on mtDNA patterns of populations. The 590-bp region of the cytochrome oxidase subunit I gene of the mitochondrial genome was sequenced for 220 specimens that were collected from 18 localities along a transect line (approximately 1400 km). Our results clearly show that the gypsy moth populations of the vast Siberian territory are not subdivided. This result can be explained by extensive genetic exchange among local populations. Taking into account that the flight ability of *L. dispar* females is low, we suggest that spreading occurs through a ballooning of early instar larvae. This hypothesis was confirmed by the correspondence of the direction of outbreak movement and dominant winds along with the observation of ballooned larvae far from a forest edge.

Introduction

The spatio-temporal distribution of animal populations is an important topic for population ecology. Insects are the most numerous animals on earth and widely distributed, while also being the main consumers of green biomass in the biosphere[1]. Many herbivorous insects can produce population outbreaks. Cyclic population dynamics and their genetic consequences have been an area of interest in ecology for many years[2].

Gypsy moth. *Lymantria dispar* L. 1758 (Lepidoptera: Erebidae) is one of the most dangerous forest pests of the Holarctic region, causing severe defoliation of over 300 plant species[3,4,5]. Outbreaking gypsy moth populations occupy vast territories that lead to significant defoliation of local forests[4,6,7]. A major question in *L. dispar* population dynamics is the regularity of outbreak appearances. The population dynamics of *L. dispar* is commonly cyclic with a period of seven to nine years and is characterised with high amplitudes of population fluctuations. Despite the analysis of long-time series data (ten times longer than the duration of one cycle), some long-term shifts in the cyclicity of outbreaks were established[8]. Usually, outbreaks of forest defoliators emerge synchronously at large spatial scales[9]. The occurrence of spatial waves of outbreaking populations over geographical space is an alternative phenomenon that has been demonstrated for certain forest defoliators[10]. For *L. dispar*, the synchronous character of outbreak appearance has been shown and mostly been studied in European and North American populations[11,12,13,14]. However, our observation of West Siberian (Asia) populations of *L. dispar* indicates the asynchronous character of outbreaks' appearance and sometimes the directed movement of the outbreak front (i.e., the travelling wave) in the north-east direction (see the Results section of this article) for some regions. The speed of this movement is greater than the potential ability of female moths to engage in flight. The reason for the outbreak front movement of *L. dispar* is still unclear. The spreading of *L. dispar* female moths cannot solely explain outbreak movement in West Siberia. West Siberia is a vast area of *L. dispar*'s range with a southern area comparable to the forest-steppe zone characterised by similar environmental conditions. The subspecies *L. dispar asisatica*, of which females possess a flight ability, inhabits this area[15]. Several studies have sought to estimate the maximal distance of *L. dispar* female

spreading[16,17,18,19,20]. In all of them, the maximum was estimated at 10 km/season or less. On the other hand, we observed the movement of pest outbreaks in West Siberia with much higher speed - 100-200 km/season, which means there are additional reasons for outbreak movements.

In the present work, we aimed to explain the outbreak movement phenomenon by the analysis of mtDNA patterns of populations collected from the vast territory of the West Siberian Plain in 2015-2016. In particular, we investigated the mtDNA patterns of *L. dispar* populations: *i*) to determine boundaries of populations of Western Siberia; and *ii*) to establish the effect of the travelling wave on mtDNA patterns of the moth populations. There has been extensive collection of mtDNA data from *L. dispar* populations[21,22,23,24,25]. This allows us to compare our dataset with the genetic variation in mtDNA data of similar reliably isolated areas of Europe.

Methods

Insect collection. Insects were collected at the pupae or adult stages in July 2015-2016 in the West Siberian plain (Fig 1). The sampling transect line was approximately 1400 km and included 18 localities. We collected several individuals from each locality, meaning that individuals from the same locality could be the progeny of the same females. However, all 18 localities were separated from each from other by no less than 10 km in accordance with[16,17,18,19,20] such that individuals from different localities were not the progeny of the same females. We characterised each locality in terms of population cycle phase (Table 1), and based on phase, the transect was subdivided into six areas, referred to as 'populations' with a distance between nearby populations of 150-200 km. The characterisation of *L. dispar* populations in each locality (forest stand) was

carried out based on the following criteria: i) the ratio of previous/current year egg masses (in West Siberia, this is easy because females of this region lay egg masses on the base of the tree stem to be covered by a snow layer); ii) current season defoliation level; iii) the size of female pupae/adults; iv) amount of parasitized larvae/pupae (cocoons of parasitic wasps around larvae or pupae with typical holes from flies). Thus, if we registered that the number of new eggs masses was higher than old eggs masses, but no greater than one new egg mass/tree, pupae were large enough (heavier than 1.5 g) and parasitism levels were low such that we assigned the "rising" population term. If we registered that the number of new egg masses was less than old eggs masses, the size of female pupae/exuvium was small (less than 1 g or equal sizes for exuvium) and parasite abundance was high, we assigned the "decline" population term. If we registered severe defoliation of birch stands, which assists with extremely high numbers of egg masses (10-40/tree), we employed the "peak" population term. The "troughs" population term was assigned when extremely low density (less than 0.005 eggs mass/tree) existed.

Fig 1. Localities where specimens of *Lymantria dispar* were collected.

Table 1. Localities of West Siberia where *L. dispar* specimens were collected.

Population name	Locality name and year of collection (internal number*)	Phase of population cycle	n	coordinates
Chulym	Shaidurovo, 2015 (14)	Rising	2	N54.29 E81.15
Chulym	Shaidurovo, 2016 (14)	Rising	10	N54.29 E81.15
Chulym	Bazovo, 2015 (15)	Rising	14	N54.34 E81.13
Chulym	Bazovo, 2016 (15)	Rising	12	N54.34 E81.13

Chulym	Noname, 2015 (16)	Rising	5	N54.55 E80.52
Chulym	Noname, 2016 (16)	Rising	7	N54.55 E80.52
Chany	Starye Karach, 2015 (1)	Rising	4	N55.28 E77.02
Chany	Noname, 2015 (2)	Rising	1	N55.30 E77.10
Chany	Chany, 2015 (3)	Rising	7	N55.24 E76.49
Chany	Chany, 2016 (3)	Rising	9	N55.24 E76.49
Omsk	Tatarsk, 2015 (4)	Peak	11	N55.10 E75.53
Omsk	Tatarsk, 2016 (4)	Decline	12	N55.10 E75.53
Omsk	Krasny Yar, 2015 (5)	Decline	17	N55.13 E72.53
Omsk	Krasny Yar, 2016 (5)	Decline	3	N55.13 E72.53
Omsk	Lubinsky, 2015 (6)	Peak	9	N55.10 E72.43
Omsk	Lubinsky, 2016 (6)	Decline	6	N55.10 E72.43
Ishim	Novolokti, 2015 (7)	Decline	2	N56.03 E69.13
Ishim	Loktyash, 2015 (8)	Decline	1	N55.56 E68.54
Ishim	Berduzhie, 2015 (9)	Decline	5	N55.48 E68.23
Ishim	Berduzhie, 2015 (9)	Decline	5	N55.48 E68.23
Tum-Ish	Bolshoy Krasnoyar, 2015 (13)	Decline	2	N56.30 E67.45
Tumen	Kyshtyrla, 2015 (10)	Decline	1	N56.57 E65.44
Tumen	Noname, 2015 (11)	Decline	19	N56.49 E65.50
Tumen	Noname, 2016 (11)	Peak	11	N56.49 E65.50
Tumen	Kirovskiy, 2015 (12)	Decline	15	N56.42 E65.42
Tumen	Kirovskiy, 2016 (12)	Peak	18	N56.42 E65.42
Trans-Ural	Kamensk-Uralsky, 2015 (19)-	Troughs	5	N56.47 E61.73
Trans-Ural	Chebarkul, 2016 (22)	Local small rising with	7	N54.75 E60.30

		decline in following years		

In the outbreaking populations, insects were collected by net (males) or by hand (females). For low population densities (i.e., trough phase), adult gypsy moths were caught via pheromones or light lures. Adults that were caught were stored in 95% ethanol. No permits for a field collection were required for this study, since the national forests in Russia are freely accessible. No protected species were sampled.

Spatial-temporal distribution of *L. dispar* outbreaks. To determine the spatial-temporal distributions of pest outbreaks, we considered time-series data of gypsy moth densities, measured as the square of defoliated forests per year in the Sverdlovsk, Chelyabinsk, Tyumen, Kurgan, Omsk, and Novosibirsk oblasts (Fig 2). The measurement of defoliation level was conducted in accordance with instructions recommended by the Federal Agency of Forest Services[26]. For calculation, we included the areas of forests where average defoliation levels exceeded 25%[26]. We utilised spectral analysis to calculate the cyclicity of outbreaks, a basic method for assessing cyclic oscillations across a time series[27,28,29]. The presence of a peak in the spectral density indicates the existence of cyclic components in a time series. However, it is not possible to correctly calculate the spectral density function for any time series. The time series studied should be stationary and its mean values and standard deviation should not change over time[27]. Otherwise, if there is a trend, the spectrum will be distorted.

Fig 2. Time-series data of *Lymantria dispar* outbreak areas in oblasts of the Russian Federation in West Siberia.

To reduce the dispersion of the studied time series and switch over to the logarithmic scale, all values of the defoliation areas, $x\ (i)$, were replaced by $x´ = x + 1$, which permitted us to transform the data with zero values of defoliation areas correctly: if $x=0$, then $ln\ x = -\infty$ but $ln\ x´ = ln\ (0 + 1) = 0$.

Further, with the analysed time series, it was necessary to "clear" high-frequency noise. The Hunn filter was applied for this purpose[30]:

$$L(x(i))=0.24x(i-1)+0.52x(i)+0.24x(i+1).$$

The carried out transformation of the time series allows tone to work correctly with the data in the program, Statistica 10.0.

To assess the coherence of time series in population dynamics, the cross-correlation function is used. The cross-correlation function, $p_{xy}\ (k)$, can be calculated for two stationary time series, x and y, with mean values, μ_x and μ_y, and standard deviations, σ_x and σ_y [27]:

$$P_{xy}(k)=(E|(x(t-k)-\mu_x)*(y(t)-\mu_y)|)/\ \sigma_x\ \sigma_y$$

where E is the operator of the mathematical expectation and $k = 0. \pm 1. \pm 2$ is the time delay.

The absence of a time delay indicates that the time series are synchronous such that $k = 0$ and the value $p_{xy}\ (0) \rightarrow 1$. The time series are coherent with a delay equal to the value of k if the maximum of the cross-correlation function falls on the value, $k = \pm 1$, $k = \pm 2$, etc. and the value of the cross-correlation function, $p_{xy}\ (k) \rightarrow 0$, for any values of k for non-conjugate time series[27].

For a more visible demonstration of outbreaks moving, we used the detailed data of the vicinities of the Novosibirsk oblast for generation of a movie (see Supplementary Video and Figure S1). The outbreaks square data were provided by the Forest Agency of Russia and the Novosibirsk branch of the Forest Protection Service.

DNA extraction, amplification, and sequencing. Total DNA was extracted from a leg of every specimen by incubation of homogenate in digestion buffer (see[31]). The 590-bp region of the cytochrome oxidase subunit I (*COI*) gene was sequenced for 220 specimens. The first part of the mtDNA amplicons (46 samples) were produced with a primer set, LepF1/LepR1[32], according to the original protocol. A second part (174 samples) was created with the primer set specific to the *L. dispar* mitochondrion genome, LepF2 5′-TACCGCTTAAACTCAGCCAT-3′and LepR2 5′-GAGGTAAAGTAAGCTCGTGT-3′, which allowed a more effective acquisition of amplicons. The LepF1/R1 primer set produced an amplicon of the 1511-2168 mtDNA region according to GenBank Accession No. FJ617240 and LepF2/R2 produced a 1457-2363 region. PCRs were carried out in a 30-μL volume with 'BioMaster HS-Taq PCR (2x)' (BioLabMix, Novosibirsk, Russia) with PCR cycling at 95°C for 5 min, 35 cycles at 95°C for 15 s, 55°C for 30 s, 72°C for 1 min, and final elongation at 72°C for 5 min. The amplicons were purified via the ZymocleanTM Gel DNA Recovery Kit (Zymo Research. USA) according to the manufacturer's instructions and sequenced with an automatic capillary sequencer with PCR primers under the BigDye® v. 3.1 (Applied Biosystems) protocol.

Data analysis. We utilised mtDNA marker COI to determine boundaries of *L. dispar* populations of Western Siberia. Obviously, for mtDNA markers, these boundaries must be determined by maternal inheritance. As the females of *L. dispar* are relatively poor fliers (flight ability is two orders less than considered area) and the local populations regularly suffer dramatic changes in effective population size, genetic drift should strongly influence mtDNA variation (content and frequencies). So, we assumed that local populations would drastically differ from each other with respect to mtDNA inheritance owing to genetic drift. We neglected the effect of other evolution factors. Selection has no significant effect and deleterious variants were quickly eliminated. It should be noted beneficial mutations in mtDNA is just theoretical. In terms of moth spreading/gene flow, the flight ability of *L. dispar* females is insufficient to cover the scale of the studied area. Mutations have too low a rate for our case. Nucleotide sequences of the *COI* gene were deposited in GenBank under accession numbers MK041668 - MK041887. The dataset of Europe populations was retrieved from BOLD Systems (www.boldsystems.org) to compare *L. dispar* mtDNA diversity of the European (see accession numbers in text S1) and Siberian regions. We made use of the European area for several reasons: *i*) the comparable scale of an uninterrupted range of *L. dispar*; *ii*) many mtDNA characterized specimens (Figure S2); *iii*) there was no effect of strong inbreeding on the population to compare with North American populations of *L. dispar* (i.e.[23]); and *iv*) populations of Europe and West Siberia were reliably isolated. The alignments of nucleotide sequences were generated by the MUSCLE programme[33] that was integrated into Mega6 software[34]. DNA polymorphism: number of polymorphic sites (S), number of haplotypes (h), haplotype diversity (Hd), nucleotide diversity (Pi); and population analysis: values of Tajima D, Fu's Fs, and F_{st} were performed using DnaSP v5[35]. A TCS gene network[36] was

performed by PopArt[37] to represent genealogical relationships among haplotypes and those frequencies.

Results

Spatial-temporal distribution of *L. dispar* outbreaks.

A quarter of a century time series data of *L. dispar* outbreaks collected in the West Siberia territory (distance between extremes localities is roughly 1400 km) demonstrated that there is a cyclic component in a temporal context (Fig 3). In particular, the peaks of the spectral power for Sverdlovsk, Chelyabinsk, Tyumen, and Kurgan regions were the same and the frequency was f_{max}=0.045 1/year, meaning that the between-peak period (L = $1/f_{max}$) was 22 years. For the Novosibirsk and Omsk oblasts, cycles were two-fold more regular (Table 2).

Fig 3. Spectral densities of log-transformed time series of squares of defoliated areas (ha): 1 - Sverdlovsk oblast, 2 - Chelyabinsk oblast, 3 – Tyumen oblast, 4 – Kurgan oblast, 5 – Omsk oblast, 6 – Novosibirsk oblast.

Table 2. The characteristics of spectral density for the time series of defoliated forests (measured as hectares of defoliated area) after *L. dispar* outbreaks occurred in West Siberia.

oblasts	Frequency f_{max} (1 x $year^{-1}$) of maximum of spectral density	Value of peak of spectral density maximum	Cyclicity of outbreaks L = $1/f_{max}$. years
Sverdlovsk	0.045	34.7	22.2

Chelyabinsk	0.045	103.1	22.2
Tyumen	0.045	74.6	22.2
Kurgan	0.045	107.5	22.2
Omsk	0.09	106.5	11.1
Novosibirsk	0.09	69.8	11.1

No strict synchrony for outbreaks in the studied areas was observed (Fig 2). The statistical analysis of the cross-correlation function shows that outbreaks were mostly coherent between different regions, and this indicated the temporal delay between comparing areas (Table 3). It was difficult to ascertain the particular direction of a spatial-temporal distribution of outbreaks when analysis was carried out for a high-scale area. It was easier to determine when outbreaks were analysed in a lesser scale, for example, within the Novosibirsk region (Table S2). For example, the movie (Supplementary Video, Figure S1) and delay value "k" in Figure 4 both demonstrate the temporal delay of outbreak movement in the spatial context in the north-east direction. Thus, we provided evidence for the travelling wave phenomenon for outbreaking populations of *L. dispar,* which concurred with the direction of dominant winds during the spring period in that area.

Table 3. The characteristics of cross-correlation functions between defoliated areas of comparing oblasts of the Russian Federation; the values of function/t-value are above the main diagonal and the temporal delays (k, years) are below.

	Sverdlovsk	Chelyabinsk	Tyumen	Kurgan	Omsk	Novosibirsk

Sverdlovsk		0.69/3.17*	0.60/2.5*	0.59/2.42*	0.21/0.98	0.66/2.95*
Chelyabinsk	-2		0.66/2.95*	0.65/2.94*	0.35/1.32*	0.56/2.17*
Tyumen	-6	-3		0.97/4.65*	0.51/2.43*	0.77/3.17*
Kurgan	-6	-3	0		0.49/1.89*	0.83/3.42*
Omsk	n.s.	-9	0	-8		0.75/3.53*
Novosibirsk	5	-8	-6	-6	1	

Genetic diversity of Siberian populations. A 590-bp fragment of the *COI* gene was sequenced for 220 gypsy moth individuals collected in 18 localities of West Siberia and Trans-Ural in the 2015-2016 seasons. Genetic diversity of the studied populations was rather low. Only 14 polymorphic sites (S) were detected. Nucleotide diversity (Pi) was 0.00156 and the average genetic distance among individuals was 0.0016. Sixteen haplotypes (h) were observed with a haplotype diversity (Hd) of 0.669. Three haplotypes (I, II, and III) made up 93% of the total sample (Fig 5, Table S1). Notably, these main haplotypes differed by one to two mutations (Fig 5). Haplotypes I and II were also found in other regions of the *L. dispar* range, including North America, Europe, and Asia[21,22,23,24,38], whereas haplotype III was unique. Negative values of Tajima's D (-1.50781. $p>0.10$) and Fu's Fs (-10.503. $p=0.0$) indicated population expansion[39,40]. The star-like net topology (Figure 5) was consistent with population expansion[41,42]. The pairwise F_{st} values indicated that maternal inheritance of populations was very closely related (Table 4). Hence, genetic data of maternal inheritance indicated that all *L. dispar* populations of the West Siberian plain could be considered large non-subdivided populations.

Fig 5. TSC network of mitochondrial haplotypes of West Siberia – Ural *L. dispar* populations.

Table 4. Pairwise F_{st} distance between populations of *L. dispar*.

Population (number of samples)	Chulym	Chany	Omsk	Ishim	Tyumen
Chulym (50)	x				
Chany (21)	0.0028	x			
Omsk (58)	-0.0034	0.0357	x		
Ishim (15)	-0.0112	0.0566	-0.0155	x	
Tumen (64)	-0.0064	0.0459	0.0031	-0.0312	x
Ural (12)	0.0238	-0.0257	0.0292	0.0899	0.0779

To compare this phenomenon of low genetic variation of *L. dispar* populations in the West Siberian plain territory with other parts of the herbivore's range, we used the data of European *L. dispar* populations. Although the European population samples were three-fold less than West Siberian samples, the genetic diversity of European populations was higher. In particular, S=16, Pi=0.00278, h=16, Hd=0.697, and the F_{st} between European and West Siberian populations was 0.08241. In European samples, we found that haplotype I was a major variant, haplotype II a minor variant, a few samples were the haplotype IV and there were many unique minor variants. Qualitative differences in

haplotype content between these areas were exemplified by the TSC net (Fig 6). Although limitations of the data did not allow European *L. dispar* to be divided into different populations, the analysis indicated a noticeably larger diversity in European populations than in West Siberia.

Fig 6. TSC network of relationships of mtDNA haplotypes of West Siberian and European *L. dispar* populations.

Discussion

Our time series results for defoliated areas clearly show that for the studied area, the duration of population cycle is approximately 11 years and for some areas, one peak was missed and outbreaks occurred every 22 years. This twice-as-long cyclicity is related with methodical lack because time series lines are restricted by only 25 variables, and the range of *L. dispar* does not well overlap with administrative division of Russian oblasts. Cross-correlation analysis readily demonstrates that most regions are characterised by the coherent trait of outbreak distributions in a spatial-temporal context (see Table 3). When we assessed the distribution of outbreaks at a small scale where the *L. dispar* range was well overlapped with administrative regions (like the Novosibirsk oblast), we registered directed movement of outbreaking areas in the north-east direction (Fig 4, Supplementary Video). Same-year synchrony was registered mostly for bordering areas (Table S2). Hence, the analysis of both large and small scales demonstrate that outbreaks taking place in West Siberia are mostly coherent, i.e., in the same period of time in different spots of space, we will find different phases of the population cycle of *L. dispar*.

Fig 4. Cross-correlation function between time series data of forest defoliation (ha) for Novosibirsk oblast indicating the temporal delay in the north-east direction: for Karasuk and Kupino districts (a), Karasuk and Krasnozerskiy districts (b), and Karasuk and Kujbyshev districts (c).

Our mtDNA data clearly shows that the diversity of the West Siberian populations of *L. dispar* was low; there were no noticeable differences among the West Siberian plain populations geographically in terms of different gypsy moth densities of population. Moreover, these variants were the same or closely related mtDNA haplotypes that were located in Europe and North America. Therefore, we concluded that the studied Siberian populations are non-subdivided in terms of mitochondrial inheritance. A non-subdivided pattern and low mtDNA diversity for so vast a territory implies extensive genetic exchange between local populations. In contrast, the genetic diversity of European *L. dispar* populations was higher in comparison with West Siberian populations. This phenomenon could be explained by isolation of populations affected by anthropogenic factors, namely pest control and others associated with the effects of urban territories in Europe. It is known that transport traffic could be also involved in the spreading of *L. dispar*[43,44,45]. However, we assume that transport is not heavily involved in the spreading of Siberian females/egg masses because: i) the traffic intensity and net of roads in Siberia is much less than in Europe (while mtDNA diversity is higher in Europe); and ii) Siberian egg masses of *L. dispar* diapause under snow layers, laying eggs near the ground for successful overwintering (Figure S3), i.e., eggs laid on cargos and vehicles will not be covered by snow. Thus, the role of transport in successful spreading of *L. dispar* in Siberia seems to be low.

The uniform mtDNA structure of West Siberian gypsy moth populations could be explained by broad spreading of females. As we mentioned in the Introduction, the flying ability of *L. dispar* females is not advanced enough to provide so low a mtDNA diversity with so great a distance. The best gypsy moth flyers inhabit Far East populations, where their maximum activity is estimated from approximately 1-2 km[18] to 10 km[19] per season. Rozkhov and Vasilyeva documented the extremely sophisticated flight abilities of *L. dispar* females in Siberian populations[46]. However, our observations over a 20-year period (Martemyanov. personal observation) and data published earlier[16] do not confirm this observation. Therefore, the spreading of females would be insufficient to explain the low diversity of such a huge region of West Siberia. We suggest that the spread of *L. dispar* results mainly from the ballooning of small instar larvae on threads that are often used by other Lepidoptera species. This explanation for no differences in mtDNA structure for two temporally distinct outbreaking populations of *Malacosoma californicum* was also provided by Franklin et al.[47].

According to a review by Bell et al.[48], the spreading distance of lepidopteran larvae by ballooning does not exceed several kilometres. In particular, the ballooning distance of European populations of *L. dispar* was directly estimated by net trapping and did not exceed 1 km[43,49]. Yet, in West Siberia, hundreds of ballooned *L. dispar* larvae were found to be attached to electricity support poles, which were as far as 15 km from the nearest forest edge (Bakhvalov. personal communication). The same scale of ballooning distance was also indirectly shown for the winter moth, *Operophtera brumata* L[50], when researchers studied the genetic structure of populations, while an earlier investigation noted much shorter ballooning distances[51].

It is significant that the spatial-temporal distribution of *L. dispar* outbreaks for certain regions of West Siberia occurred as a travelling wave in a north and north-east direction, as recorded over the past quarter of a century (Supplementary Video, Fig 4). This direction is in line with the direction of the dominant wind in this area in spring[52], which is the period of larval hatching in West Siberia[53,54] and follows ballooning, indirectly reflecting the importance of ballooning in open areas, such as the forest step zone or archipelagos.

We can conclude that the vast territory (over 1000 km) with similar climatic conditions (continental climate) and the same landscape (plain of forest-step zone) is inhabited by an non-subdivided population of *L. dispar*, which maintains its mtDNA structure independent of population-cycle phases. Although there is genetic drift during the trough phase, the mtDNA structure remains stable (recovers) in the following years. The following facts indicate the major role of long spreading by ballooning: *i*) the genetic similarity of mtDNA patterns of low mobility species in the vast territory; *ii*) the correspondence of the direction of outbreak movement with dominant winds; and *iii*) the observation of ballooned larvae far from a forest edge. We assume that there is potential for Lepidoptera larvae ballooning being underestimated to date. The flight ability of females is also implicated in the low mtDNA diversity of Siberian populations. However, this factor seems to operate mostly at the local scale.

Acknowledgements

We would like to thank Dr. Vasiliy Ponomarev for providing specimens of *L. dispar* from the Yekaterinburg oblast. We also thank Nikita Kasianov who helped in the

collection of samples. In addition, we express gratitude to Andrey Broshkov and Vladimir Judkin for helping with the idea and preparation of illustrations.

References

1. Schoonhoven LM., van Loon JJA & Dicke M. Insect-Plant Biology. 2nd ed. Oxford: University Press; 2005.
2. Elton CS. Periodic fluctuations in the numbers of animals: their causes and effects. British J Exp Biol. 1924; 2:119-163.
3. Giese RL & Schneider ML. Cartographic comparisons of Eurasian gypsy moth distribution (*Lymantria dispar* L.; Lepidoptera: Lymantriidae). Entomol News. 1979; 90:1-16.
4. Bakhvalov SA, Koltunov EV & Martemyanov VV. Factors and ecological mechanisms of population dynamics of forest insects-phyllophages. Russia. Novosibirsk. SB RAS Press. pp. 296; 2010.
5. Ponomarev VI, Ilinykh AV, Gninenko, JI, Sokolov GI & Andreeva EM. Gypsy moth in Trans-Ural and Western Siberia. 1st ed. Russia. Yekaterinburg. Russia. Ural branch of Russian academy of science; 2012.
6. Il'ynskiy AI & Tropin IV. The control. calculation. and prediction of outbreaking forest folivorous insects in the forests of USSR. 1st ed. Russia. Moscow. Forest industry publisher; 1965.
7. Qian TR. Gypsy moths' damage to America and measures taken by United States Department of Agriculture. Plants Quarantine. 14, 317-318; 2000.
8. Allstadt AJ, Haynes KJ, Liebhold AM & Johnson DM. Long-term shifts in the cyclicity of outbreaks of a forest-defoliating insect. Oecologia. 2013; 172:141-51.
9. Myers JH & Cory JS. Population Cycles in Forest Lepidoptera Revisited. Annu Rev Ecol Evol Syst. 2013; 44:565-592.

10. Tenow O, Nilssen AC, Bylund H, Pettersson R, Battisti A, Bohn U *et al*. Geometrid outbreak waves travel across Europe. J Anim Ecol. 2012; 82:84-95.
11. Liebhold A, Elkinton J, Williams D & Muzika R. What causes outbreaks of the gypsy moth in North America? Popul Ecol. 2000; 42:257-266.
12. Soukhovolsky VG, Ponomarev VI, Sokolov GI, Tarasova OV & Krasnoperova PA. Gypsy moth *Lymantria dispar* L. in the South Urals: Patterns in population dynamics and modeling. Zh Obshch Biol. 2015; 76:179-194.
13. Myers JH. Synchrony in outbreaks of forest lepidoptera: a possible example of the Moran effect. Ecology. 1998; 79:1111-1117.
14. Johnson DM, Liebhold AM, Bjornstad ON & Mcmanus ML. Circumpolar variation in periodicity and synchrony among gypsy moth populations. J Anim Ecol. 2005; 74:882-892.
15. Pogue MG & Schaefer PW. A review of selected species of Lymantria Hübner (1819) (Lepidoptera: Noctuidae: Lymantriinae) from subtropical and temperate regions of Asia, including the descriptions of three new species, some potentially invasive to North America. Washington; 2007.
16. Gninenko YuI. Geographicheskie formy neparnogo shelkopryada v Severnoy i Centralnoy Asii. Lesnoy vestnik MGUL. 2003; 2:166-174.
17. Liebhold AM, Turcáni M & Kamata N. Inference of adult female dispersal from the distribution of gypsy moth egg masses in a Japanese city. Agric Forest Entomol. 2008; 10:69–73.
18. Iwaizumi R, Arakawa K & Koshio C. Nocturnal flight activities of the female Asian gypsy moth. *Lymantria dispar* (Linnaeus) (Lepidoptera: Lymantriidae). Appl Entomol Zool. 2010; 45:121-128.
19. Yang F, Lou, Y & Shi J. The influence of geographic population, age, and mating status on the flight activity of the Asian gypsy moth *Lymantria dispar* (Lepidoptera: Erebidae) in China." App Entomol Zool. 2017; 52:265-270.

20. Baranchikov YN. Ecological basis of the evolution of host relationships in Eurasian gypsy moth populations (ed. Wallner, W. E. & McManus, K. C.). In: Proceedings, Lymantriidae: A Comparison of Features of New and Old World Tussock Moths, 319-338. Broomall, 1988.
21. Bogdanowicz SM, Schaefer PW & Harrison RG. Mitochondrial DNA variation among worldwide populations of gypsy moths. *Lymantria dispar*. Mol Phylogenet Evol. 2000; 15:487-495.
22. deWaard JR, Mitchell A, Keena MA, Gopurenko D, Boykin LM, Armstrong KF e*t al*. Towards a global barcode library for *Lymantria* (Lepidoptera: Lymantriinae) tussock moths of biosecurity concern. PLoS ONE. 2010; **5:**12, https://doi.org/10.1371/journal.pone.0014280.
23. Wu Y, Molongoski JJ, Winograd DF, Bogdanowicz SM, Louyakis AS, Lance DR e*t al*. Genetic structure. admixture and invasion success in a Holarctic defoliator. the gypsy moth (*Lymantria dispar*. Lepidoptera: Erebidae). Mol Ecol. 2015; 24:1275-1291.
24. Djoumad A, Nisole A, Zahiri R, Freschi L, Picq S, Gundersen-Rindal DE *et al*. Comparative analysis of mitochondrial genomes of geographic variants of the gypsy moth. *Lymantria dispar*. reveals a previously undescribed genotypic entity. Sci Rep. 2018; 7:14245.
25. Inoue MN, Suzuki-Ohno Y, Haga Y, Arai H, Sano T. Martemyanov VV *et al*. Population dynamics and geographical distribution of the gypsy moth, *Lymantria dispar*, in Japan. Forest Ecol Management. 2019; 434:154–164.
26. Rukovodstvo po planirovaniju, organizacii i vedeniju lesopatologicheskogo obsledovanija (The instruction for planning, organizing and realizing of observation for forests diseases and pests). Supplementary to order #527 from 29.21.2007, Federal Agency of Forest Services.
27. Box GEP & Jenkins GM. Time series analysis: forecasting and control. San Francisco: Holden-Day; 1970.

28. Kendall MG & Stewart A. The advanced theory of statistics: design and analysis, and time series. L.: C. Griffin. Hafner; 1976.
29. Marple SL, Jr. Digital spectral analysis: with applications. Englewood Cliffs, N.J.: Prentice-Hall; 1987.
30. Hamming RW. Digital filters. Englewood Cliffs, N.J.: Prentice Hall; 1989.
31. Ilinsky YY, Tokarev YS, Bykov RA, Yudina MA, Pavlushin SV, Inoue MN *et al*. Detection of bacterial symbionts (Wolbachia. Spiroplasma) and eukaryotic pathogen (Microsporidia) in Japanese populations of gypsy moth species (Lymantria spp.). Evraziat Entomol Z. 2017; 16:1-5.
32. Hebert PDN, Penton EH, Burns JM, Janzen DH & Hallwachs W. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator. Proc Natl Acad Sci U.S.A.2004; 101:14812-14817.
33. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004; 32:1792-1797.
34. Tamura K, Stecher G, Peterson D, Filipski A & Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Boil.Evol. 2013; 30:2725-2729.
35. Librado P & Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009; 25:1451-1452.
36. Clement M, Posada D & Crandall KA. TCS: a computer program to estimate gene genealogies. Mol Ecol. 2000; 9:1657-1659.
37. Leigh JW & Bryant D. PopART: full-feature software for haplotype network construction. Methods Ecol Evol. 2015; 6:1110-1116.
38. Picq S, Keena M, Havill N, Stewart D, Pouliot E, Boyle B *et al*. Assessing the potential of genotyping-by-sequencing-derived single nucleotide polymorphisms to identify the geographic origins of intercepted gypsy moth (*Lymantria dispar*) specimens: A proof-of-concept study. Evol Appl. 2018; 11:325-339.

476 39. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA
477 polymorphism. Genetics. 1989; 123:585-595.
478 40. Fu YX. Statistical tests of neutrality of mutations against population growth. hitchhiking and
479 background selection. Genetics. 1997; 147:915-925.
480 41. Ferreri M, Qu W & Han B. Phylogenetic networks: a tool to display character conflict and
481 demographic history. African J Biotech. 2011; 10:12799-12803.
482 42. Bandelt HJ, Forster P, Sykes BC & Richards MB. Mitochondrial portraits of human
483 populations using median networks. Genetics. 1995; 141(2):743-753.
484 43. Doane CC & McManus ME (Eds.). The Gypsy Moth: Research toward Integrated Pest
485 Management. Washington; 1981.
486 44. Hajek AE & Tobin PC. North American eradications of Asian and European gypsy moth
487 (eds. Hajek AE, Glare TR. & O'Callaghan M) In: Use of Microbes for Control and
488 Eradication of Invasive Arthropods. New York; 2009.
489 45. Liebhold AM & Tobin PC. Growth of newly established alien populations: comparison of
490 North American gypsy moth colonies with invasion theory. Popul Ecol. 2006; 48:253-262.
491 46. Rozkhov AS & Vasilyeva TG. Gypsy moth in Central and Eastern Siberia in *Neparny*
492 *Shelkoprayo v srednai I Vostochnoi Sibiti* (Russian Academy of Sciences. Siberian Branch)
493 4-19. Novosibirsk, Russia; 1982.
494 47. Franklin MT, Myers JH & Cory JS. Genetic Similarity of Island Populations of Tent
495 Caterpillars during Successive Outbreaks. 2014; PLoS ONE. 9. e96679;
496 10.1371/journal.pone.0096679.
497 48. Bell JR, Bohan DA, Shaw EM, Weyman, GS. Ballooning dispersal using silk: world fauna.
498 phylogenies. genetics and models. Bull Entomol Res. 2005; 95:69-114.
499 49. Weseloh RM. Evidence for limited dispersal of larval gypsy moth. *Lymantria dispar* L.
500 (Lepidoptera: Lymantriidae). Can Entomol. 1997; 129:355-361.

50. Leggett HC. et al. Population genetic structure of the winter moth. *Operophtera brumata* Linnaeus. in the Orkney Isles suggests long distance dispersal. Ecol Entomol. 2011; 36:318-325.
51. Edland T. Wind dispersal of the winter moth larvae *Operophtera-brumata* Lepidoptera Geometridae and its relevance to control measures. Norsk Entomologisk Tidsskrift. 1971; 18:103-107.
52. Kravcov VM & Donukalova RP. *Geography of Novosibirsk oblast (Geographia Novosibirskoi oblasti)*. INFOLIO press; 1999.
53. Martemyanov VV, Pavlushin SV, Dubovskiy IM, Yushkova YY, MorosovSV, Chernyak EI *et al*. Asynchrony between host plant and insects defoliator within a tritrophic system: The role of herbivore innate immunity. PLoS ONE. 2015; 10: e0130988. doi:10.1371/journal.pone.0130988.
54. Martemyanov VV, Belousova IA, Pavlushin SV, Dubovskiy IM, Ershov NI, Alikina TY *et al*. Phenological asynchrony between host plant and gypsy moth reduces insect gut microbiota and susceptibility to *Bacillus thuringiensis*. Ecol Evol. 2016; 6:7298-7310.

Supplementary information

Supplementary video