MOLECULAR ECOLOGY

Molecular Ecology (2014) doi: 10.1111/mec.12913

# Temporal relationship between genetic and warning signal variation in the aposematic wood tiger moth (*Parasemia plantaginis*)

JUAN A. GALARZA, OSSI NOKELAINEN, ROGHAEIH ASHRAFI, ROBERT H. HEGNA and JOHANNA MAPPES

*Centre of Excellence in Biological Interactions, Department of Biological and Environmental Science, University of Jyväskylä Survontie 9, Jyväskylä 40500, Finland*

## Abstract

**Many plants and animals advertise unpalatability through warning signals in the form of colour and shape. Variation in warning signals within local populations is not expected because they are subject to directional selection. However, mounting evidence of warning signal variation within local populations suggests that other selective forces may be acting. Moreover, different selective pressures may act on the individual components of a warning signal. At present, we have a limited understanding about how multiple selection processes operate simultaneously on warning signal components, and even less about their temporal and spatial dynamics. Here, we examined temporal variation of several wing warning signal components (colour, UV-reflectance, signal size and pattern) of two co-occurring colour morphs of the aposematic wood tiger moth (*Parasemia plantaginis*). Sampling was carried out in four geographical regions over three consecutive years. We also evaluated each morph's temporal genetic structure by analysing mitochondrial sequence data and nuclear microsatellite markers. Our results revealed temporal differences between the morphs for most signal components measured. Moreover, variation occurred differently in the fore- and hindwings. We found no differences in the genetic structure between the morphs within years and regions, suggesting single local populations. However, local genetic structure fluctuated temporally. Negative correlations were found between variation produced by neutrally evolving genetic markers and those of the different signal components, indicating a non-neutral evolution for most warning signal components. Taken together, our results suggest that differential selection on warning signal components and fluctuating population structure can be one explanation for the maintenance of warning signal variation in this aposematic species.**

*Keywords*: colour polymorphism, Erebidae, fluctuating populations, melanization, UV signalling

*Received 13 November 2013; revision received 5 September 2014; accepted 8 September 2014*

## Introduction

Analyses of temporal variation in phenotype and genotype frequencies can provide valuable insights into selective processes affecting natural populations. However, studies reporting temporal replicates comprise only a small fraction of the literature (Siepielski *et al.* 2009). Moreover, temporal studies often focus on changes in either phenotypes or genotypes (i.e. Grant & Grant 2002; Pilot *et al.* 2010). Simultaneous analyses of phenotypes and genotypes are thus essential to understand the temporal strength and form of selection. This is particularly true for seasonal populations whose fitness depends on the relative frequency of their phenotypes (Hendry & Day 2005). For instance, aposematic

Correspondence: Juan A. Galarza, Fax: +358 (0)14 617 239; E-mail: juan.galarza@jyu.fi

species use coloration as a warning signal to advertise unpalatability to potential predators (Cott 1940; Edmunds 1974). Directional selection by predators reduces signal variation, increasing the frequency of the phenotype with the most effective warning signal (Symula *et al.* 2001; Mallet 2010). This in turn induces a reproductive bias towards the most common phenotype, which is then reflected in a genotype shift in the subsequent season.

Diversity in warning signals within the same population is not expected because less variable warning signals are easier for predators to learn and remember (Joron & Mallet 1998; Beatty *et al.* 2004; Rowland *et al.* 2007). Despite this, local variation in warning signals has been reported in different species such as frogs and newts (Mochida 2011; Rojas & Endler 2013). Variation can result from differential predation (Endler & Mappes 2004; Valkonen *et al.* 2012), resource allocation (Grill & Moore 1998; Ojala *et al.* 2007; Lindstedt *et al.* 2010b) and environmental (Davis *et al.* 2012; Hegna *et al.* 2013) or seasonal changes (Oostra *et al.* 2011). Thus, warning signals are subject to variable selection in space and time (Mappes *et al.* 2005).

While variation in warning signals has been the scope of interest of both theoreticians and empiricists (Endler & Mappes 2004; Mappes *et al.* 2005; Stevens & Ruxton 2012), little is known about selection acting on the individual components that constitute a warning signal, and even less about their temporal and spatial dynamics. The arrangement, size and distribution of warning signal components may serve different functions. For instance, warning signals of aposematic strawberry poison frogs (*Oophaga pumilio*) both serve as toxicity advertisement (Saporito *et al.* 2007; Richards-Zawacki *et al.* 2013) and are also used by females to distinguish local males, which are preferred (Maan & Cummings 2009). Similarly, some warning signals in lepidopterans are jointly involved with mate choice (Nokelainen *et al.* 2012) and with con-specific recognition where interspecies mimicry is involved (Kronforst *et al.* 2006). Furthermore, thermoregulatory needs can have an effect in warning signal composition by reducing its size, and hence limiting its efficacy (Lindstedt *et al.* 2009; Hegna *et al.* 2013). Therefore, it can be expected that different signal components experience differential selection pressures, making variation adaptive in some components and neutral in others (Ratcliffe & Nydam 2008; Wollenberg *et al.* 2008).

Given that multiple selection pressures operate simultaneously on warning signal components, it is important to study their genetic consequences in space and time. Here, we examine variation in various components of the warning signal in the wings of the aposematic wood tiger moth (*Parasemia plantaginis*) from four geographical regions (Finland, Åland, Estonia and the Alps) over three consecutive years. Two male phenotypes (hereafter morphs) occur in sympatry within these regions. The morphs are polymorphic in their warning signal coloration (yellow or white), as well as variable in their black-pattern elements in both their fore- and hindwings. Both the hindwing coloration and the amount of melanization are inherited (Lindstedt *et al.* 2009; Nokelainen *et al.* 2013), although the details of the genetic architecture of the coloration are not yet known. We examine the temporal relationship between genetic and phenotypic variation by comparing variation patterns of the warning signal components, presumably under selection, such as colour (including UV-reflectance), signal size and pattern, to those produced by neutrally evolving genetic markers (COI gene and microsatellites). We first determine whether the yellow and white morphs constitute a single genetic population within each region, and whether such a population is stable through time. If the different morphs constitute different populations at a given time, it can be seen as an indicative of assortative mating and/or disruptive selection. Second, we assess whether temporal differences exist in the warning signal components of both the fore- and hindwings for both morphs in each region. Wing sets are analysed separately because they provide a different form of protective coloration (i.e. crypsis vs. warning). Differentiation between wing sets would indicate different selection process operating simultaneously. Finally, we determine whether the pattern of temporal variation observed in each signal component can be correlated with a pattern of variation generated under neutral evolution. Negative correlations would suggest that the signal component has not evolved in a neutral fashion.

## Materials and methods

### *Study species*

The wood tiger moth (*Parasemia plantaginis*) is an aposematic moth of the Arctiidae family, widely distributed throughout the northern hemisphere (Dubatolov & de Vos 2010). This generalist species occurs in a variety of habitats, its generation time is one year (i.e. univoltinism), and overwintering occurs at the larval stage (Ojala *et al.* 2005; Lindstedt *et al.* 2010a). Throughout its range, females exhibit continuous colour variation from orange to red (Lindstedt *et al.* 2010b), with a melanization pattern fairly similar to that of males. In Europe, males have heritable hindwing colour (either white or yellow) (Nokelainen *et al.* 2012) with black pigmentation covering varying amounts of the wing's surface (Hegna *et al.* 2013). The environment experienced during development

does not seem to have an effect on adult coloration, and it has been suggested that colour polymorphism is probably the result of the interaction between two or more closely linked genes (Nokelainen *et al.* 2012). Male polymorphisms are found within populations, where two colour morphs co-occur at variable frequencies (Nokelainen *et al.* 2012). Laboratory and field studies have confirmed that hindwing coloration in this species acts as a visual warning signal to deter predators, and the yellow morph seems to be better protected than the white morph (Hegna *et al.* 2013). Tits (Paridae) and dunnocks (Prunellidae) appear to be the most important avian predators for wood tiger moths (Nokelainen *et al.* 2014). Both morphs are warningly coloured because their survival increases with conspicuousness. The yellows, however, are more conspicuous than whites against most natural backgrounds (Nokelainen *et al.* 2012).

Wing coloration can also function as a secondary sexual trait in this species (Nokelainen *et al.* 2012). Under certain circumstances, the white morph has a mating advantage. For instance, in laboratory settings, after secreting defensive fluids, the white morphs had a significantly higher mating success compared with yellows (Nokelainen *et al.* 2012). Furthermore, wings with more melanization can be beneficial for thermoregulation (Hegna *et al.* 2013), inducing a trade-off between thermoregulation and signal display. Therefore, this species provides a good opportunity to examine various selection pressures affecting warning signals and coloration in general.

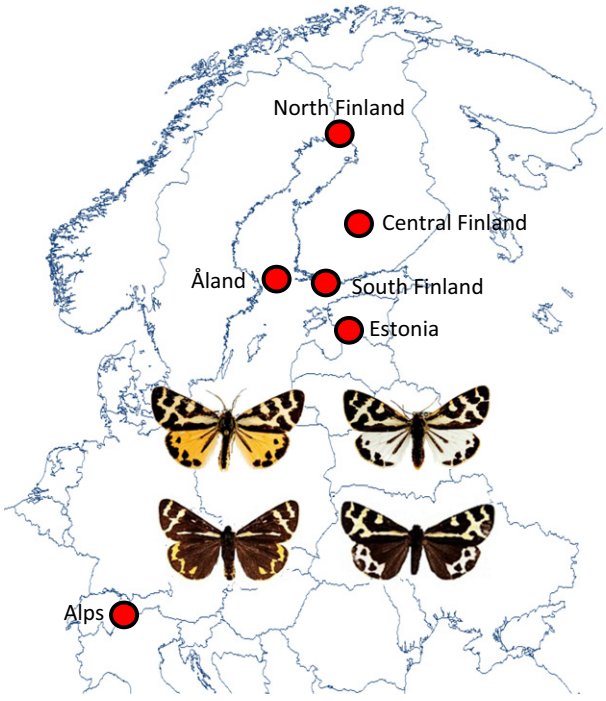


**Fig. 1** Map showing *Parasemia plantaginis* sampling locations and the two most common male morphs found in Finland (including Åland and Estonia) and in the Alpine region.

### *Sample collection*

During three consecutive summers (2009–2011), we collected a total of 537 *P. plantaginis* males throughout 58 sampling sites across Finland, the Åland Islands and Estonia. Additionally, 86 individuals were collected at three sampling sites in the Swiss and Austrian Alps during 2009 and 2010 (Fig. 1). In the Finnish and Åland regions, both morphs can be found in frequencies varying between 52% and 77% of yellows (O. Nokelainen, unpublished data). In Estonia, on the other hand, the great majority of the males are white (97%). In the Alps, both morphs show a pronounced melanization in their wing pattern. Therefore, a combined analysis in these regions provides a good opportunity to study colour polymorphisms and melanization patterns. Each sampling region had at least two sampling sites with a mean of 15.7 samples/morph/year (Table S1, Supporting information). In all regions, temporal samples were collected from the same sampling sites over two or three consecutive years. We collected individuals using butterfly nets and pheromone traps with a calling female as lure. For all samples, we employed a combination of genotypic (microsatellite, mitochondrial) and phenotypic (wing coloration – including UV-reflectance – signal size and wing pattern) analyses. All variables were analysed separately and in combination as described below.

### *Phenotypic data collection*

We used digital photography to measure the fore- and hindwing coloration of adult *P. plantaginis* males. Photographic analysis was chosen over spectrophotometry due to its high repeatability, mass scaling and easy optimization. Photographs were taken from dead individuals with an ultraviolet-sensitive FujiFilm Finepix S3 Pro UVIR digital camera under a light source emitting both visible and UV wavelengths (Arcadia Reptile D3, Salfords, UK). Two photographs were taken of each individual: (i) a regular photograph to human visual spectrum (approximately 400–700 nm) taken with a UV and an infrared blocking filter (Baader UV/IR Cut, Baader Planetarium, Mammendorf, Germany) and (ii) a UV photograph taken with a filter transmitting only the UV spectrum (approximately 300–400 nm, filter: Baader U). All specimens were photographed in one go and under the same conditions.

The saturation values were obtained from using the calibrated digital photographs (Stevens *et al.* 2007; Lindstedt *et al.* 2010b). The camera's response of RGB (red,

green and blue) channels to increasing light intensity is usually nonlinear, although direct comparison with colour saturation requires a linear response (Stevens *et al.* 2007). Therefore, we first linearized the response of each of these channels. In every photograph, a grey standard (Labsphere Spectralon diffuse saturation standard reflecting 50% of all light) was included, enabling the conversion of images to saturation data using a custom MATLAB script (see Stevens *et al.* 2007).

To analyse the variation in coloration, the pigmentation of both wings and their melanization pattern were sampled using GIMP v2.8.2 software®. From each individual, the right fore- and hindwings were measured, taking three samples from each point of interest (Fig. S1, Supporting information). We then recorded saturation values in the ultraviolet (UV approximately 300–400 nm), short (SW approximately 400–500 nm), medium (MW approximately 500–600 nm) and long (LW approximately 600–700 nm) wavelengths. The coloration of both morphs was analysed based on hue (i.e. saturation of pixels) measurements of these wavelengths by standardizing each to a proportion of the total saturation (e.g. LW/(UV+SW+MW+LW)).

The signal size (Ss) was measured as the proportion of melanization in the wing (i.e. the smaller the melanization, the larger the signal) using Paint.Net software by dividing the melanized area by the total wing area (Fig. S1, Supporting information). From the right fore- and hindwings, we measured the wing area (with the polygon selection tool), length and width (with the line selection tool). We used millimetre graph paper for scale.

To analyse the differences in the forewing pattern (Patt), we followed methods developed by Todd *et al.* (2005) with some modifications. A geometric morphometrics approach could not be implemented here because of the difficulty in identifying consistent landmarks. In particular, landmarks at black-pattern elements or at junctions between black and colour elements were not consistent in all samples. The grid approach from Todd *et al.* (2005) provides a good alternative to obtain landmark analogues, and the statistical procedures are also based on distance matrix analyses. Wings were photographed in standardized laboratory conditions as described earlier. One forewing from each specimen was traced before being copied and then pasted into Microsoft PowerPoint in Microsoft Office 2007®. The image was then adjusted on each slide to be positioned in a similar way with respect to every wing image without distorting the pattern (Fig. S3, Supporting information). Each photograph was then enlarged without distortion to fill the entire slide. Within PowerPoint, a grid measuring 35 × 35 cells was overlaid on each image. Unlike in Todd *et al.* (2005), only the pattern elements of interest were coded with a '1' in the grid. A macro was created within Microsoft Excel to fill in all other empty cells with a '0' and then to rearrange the square grid into a single column of data. Each grid then comprised a single column and was labelled with an alphanumeric code denoting population membership and other factors of interest.

### *Genetic data collection*

Total DNA was extracted from one or two moth legs using DNeasy Blood + Tissue extraction kit (Qiagen) assisted by a robot (Kingfisher). For all samples, we amplified 630 bp of the cytochrome oxidase c subunit I (COI) gene. The amplification reaction was performed in 20 μL total volume with approximately 40 ng of DNA, 2 mM of $MgCl_2$, 0.5 μM of forward and reverse primer, 0.2 mM of dNTP's, 10× Dream Taq Buffer (Fermentas) and 5 U/μL Dream Taq DNA polymerase (Fermentas). PCR conditions were as follows: an initial denaturation of 3 min at 95 °C, 30 cycles of 30 s at 95 °C, 30 s at 55 °C annealing temperature, 1.5 min at 72 °C and a final elongation step of 5 min at 72 °C. Amplification products were sequenced using the BigDye® terminator v3.1, Cycle Sequencing kit (Applied Biosystems) and ran on an ABI 3130xl Genetic Analyser (Applied Biosystems). Sequences were manually edited and aligned using SEQUENCHER version 4.9 (Gene Codes, Ann Arbor, MI) and deposited in the European Nucleotide Archive (accession no. HG764800-HG764812). The same extraction protocol was used to amplify a total of 10 species-specific nuclear microsatellite markers according to Galarza *et al.* (2011). Allele scoring of the microsatellite data was carried out using GENEMAPPER v.4.0 (Applied Biosystems).

### *Mitochondrial data analyses*

To examine the level of genetic diversity in the COI, we calculated the number of haplotypes (h), number of segregating sites (S), nucleotide diversity (pi) and haplotype diversity (H) with DNASP v. 5.10 (Librado & Rozas 2009). Haplotype relationships were analysed by constructing a network with the software TCS v.1.21 (Clement *et al.* 2000), which implements a statistical parsimony procedure, with gaps coded as a fifth character state and setting a 95% connection limit. A Tajima's D test of neutrality was implemented using DNASP v. 5.10 (Librado & Rozas 2009) to infer whether the observed haplotypes conform to a neutral model of evolution. Finally, we conducted an analysis of molecular variance (AMOVA) in Arlequin v.3.5.1.3 (Excoffier & Lischer 2010). Hierarchal comparisons were performed between years, between morphs within years and within years.

*Microsatellite data analyses*

We evaluated the temporal variability of nuclear microsatellite loci within and between morphs as follows.

*Within morphs.* To estimate the degree of temporal population structure within the morphs, we analysed allelic frequency variation at 10 species-specific microsatellite loci (Galarza *et al.* 2011). As a first step, we performed an analysis of molecular variance (AMOVA) to evaluate the distribution of genetic variation within each region and morph. Hierarchal comparisons were performed between years, between sampling sites, within years and within sampling sites. The analyses were executed in GENALEX v.6 (Peakall & Smouse 2006), and statistical significance ($P < 0.05$) was attained through 999 random permutations for each hierarchal comparison. Second, we used STRUCTURE v.2.3 (Pritchard *et al.* 2000) to infer the most likely number of clusters (populations) present in our data set. The correlated allelic frequencies model was used with no prior information included about sample origin, morph type or year of collection. We set the number of clusters (K) from one to 20 and ran 10 independent iterations consisting of three million MCMC steps each, a third of which was discarded as burnin. We then used STRUCTURE HARVERSTER (Earl & vonHoldt 2012) to assess the likelihood of the different Ks according to the Evanno *et al.* (2005) method. Finally, we used CLUMMP v.1.1.2 (Jakobsson & Rosenberg 2007) to evaluate the consistency of the results across the iterations using the full-search algorithm. The software DISTRUCT v.1.1 (Rosenberg 2004) was used to graphically display all results.

Two complementary approaches were implemented to further evaluate the degree of connectivity between regions and the temporal genetic structure within them. For these and subsequent analyses, samples were grouped by region/morph/year according to the AMOVA and STRUCTURE results. In the first approach, we tested for temporal allelic frequency differentiation ($F_{ST}$) within morphs and regions. The $F_{ST}$'s estimated probabilities were calculated using GENETIX v.4.05 (Belkhir *et al.* 1997) through 10 000 permutations according to Weir & Cockerham (1984). Bonferroni corrections were applied for accepting the alternative hypothesis of significant genetic differentiation by setting an initial significance level of 0.05 (Rice 1989). We also evaluated two complementary differentiation statistics to account for variation in sample sizes using GENODIVE v2.0 (Meirmans & Van Tienderen 2004). One statistic is based on the number of effective alleles (Jost's *D*; Jost 2008), whereas the other implements a bias correction for small sample sizes (Nei's *D*; Nei 1978b). Significance ($P < 0.05$) of both tests was attained by means of 1000 random permutations. For the second approach, we used MIGRATE v.3.2.1 (Beerli 2006) to estimate long-term connectivity between regions. The program infers migration rates through a probabilistic model based on an expansion of the coalescent theory, and thus taking into account lineage histories. The program was first run with default settings using the Bayesian inference strategy and Brownian motion as the mutational model. The output was used to parameterize three subsequent runs recording 3 000 000 genealogies, the first 450 000 of which were discarded as burnin.

Deviations from Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium within regions and years were estimated according to the level of significance determined by means of 10 000 MCMC iterations using GENEPOP v.3.4 (Raymond & Rousset 1995). Bonferroni corrections were applied for multiple comparisons setting an initial significance level of 0.05 (Rice 1989). The software package FreeNA (Chapuis & Estoup 2007) was used to test for the possibility of null alleles in each locus within regions. Genetic diversity for each morph was characterized by calculating the number of alleles per locus (Na), the effective number of alleles (independently of sample size, Ar), the fixation index ($F_{IS}$) and observed ($H_O$) and expected heterozygosities ($H_E$) computed in FSTAT (Goudet 1995) and ARLEQUIN v.3.5.1.3 (Excoffier & Lischer 2010).

*Between morphs.* To examine whether significant temporal differences exist between the morphs, we compared estimates of Ar, $H_E$ and Na for each year at each region by applying a *t*-test. We then obtained a measure of genetic distance based on the proportion of shared alleles ($D_{ps}$) using MSANALYZER v.4.05 (Dieringer & Schlötterer 2003) according to Bowcock *et al.* (1994). Here, we estimated the mean proportion of shared alleles between the morphs within sampling regions and years. We also tested for the presence of private alleles in each specific morph, that is, alleles whose frequency is $\leq 0.01$ within a morph in a given year. Finally, we performed a population assignment test between morphs within years and regions using GENALEX software v.6 (Peakall & Smouse 2006). Here, each sample was assigned either to its own population (i.e. morph) or to another population within the same year and region. For this frequency-based assignment test, we used the bias correction for population frequency (Paetkau *et al.* 2004) in which the sample to be assigned is removed from the data set before calculating the adjusted frequencies. In the case of Estonia, all comparisons were made on white morphs only because Estonian populations consist almost entirely of white morph individuals (O. Nokelainen, Personal observation). Moreover, despite intensive sampling effort in Estonia,

no samples were available for 2010 and hence, comparisons were made between 2009 and 2011.

### *Warning signal analyses*

We subdivided the warning signal of both morphs into four major components and tested for temporal differences in these between and within the morphs for both wing sets. The first component was indicated by colour (Col), composed by saturation values of UV, SW, MW and LW. The second component analysed was the UV-reflectance alone as it has been shown that in Lepidoptera, UV-reflectance can be attractive to both predators (Lyytinen *et al.* 2004) and potential mates (Brunton & Majerus 1995; Robertson & Monteiro 2005), thus being under directional and/or sexual selection. The third and fourth signal components were signal size (Ss) and forewing pattern (Patt) as described earlier.

To gain an overall perspective of how the coloration of each morph is characterized in the visible spectrum and UV, we first performed a principal component analysis (PCA) with orthomax rotation on the SW, MW, LW and UV for both the forewing and the hindwing including all samples. Second, we analysed each wavelength saturation value separately in each region to test for temporal differences within and between the morphs via analyses of variance (ANOVA) using MATLAB v.7.11 (Mathworks). This method, however, simply returns the probability that all saturation values are drawn from groups with the same mean. To test which pairs of means were significantly different, we conducted a *post hoc* analysis using the *multcompare* function in MATLAB v.7.11 (Mathworks), which implements a multiple comparison procedure using the ANOVA output. Here, pairs of means are compared correcting the significance level for multiple comparisons and returning a 95% confidence interval about the true difference between any two means. The same procedure was followed for the signal size, whereas the pattern was analysed in combination with genetic data as described below.

### *Combined genetic and phenotypic analyses*

With these combined analyses, we aimed to compare the variation pattern of the different signal components, presumably under selection, to that produced by neutrally evolving genetic markers, such as the COI fragment and our panel of nuclear microsatellites. It should be noted, however, that negative correlations cannot be taken as direct evidence of selection, but as an indication that the trait has not evolved as it would be expected under neutral evolution. To strengthen the biological meaning of the results, these analyses included only samples collected from the same sampling sites for at least two years, and from which we had complete data sets (i.e. COI, microsatellite, UV-reflectance, Col, Ss and Patt), with a mean of 31.4 samples/morph/region/year. Therefore, some comparisons were not possible due to incomplete data sets of the same samples. We analysed the different data sets simultaneously by performing a series of Mantel and partial Mantel tests, also known as casual modelling, using IBD v.1.52 (Bohonak 2002). Here, we tested for the possible correlation between a genetic distance matrix (COI, microsatellite) and a signal component distance matrix (Col, UV, Ss, Patt). For the COI, we first calculated the genetic distance between the samples according to the Jukes & Cantor (1969) evolutionary model using MEGA v.5.05 (Tamura *et al.* 2011). As no empirical substitution rates are available for our species, we chose this distance model because it assumes equal substitution rates among all segregating sites. A distance matrix was computed for the microsatellite data set in MSANALYZER v.4.05 (Dieringer & Schlötterer 2003) according to Nei's (1978a) standard genetic distance. The distance matrices for the different signal components of both the fore- and hindwings were calculated as the Euclidian distance between their respective hue values using MATLAB v.7.11 (Mathworks). In the case of the signal component 'colour', we calculated the Euclidean distance using the first principal component values of the PCA. A third matrix corresponding to the Euclidean distance of the forewing black pattern was further included in a partial Mantel test. This test determines whether the wing's black pattern has an effect on the correlation between the genetic and signal component matrices ($r$). In the same way, we tested whether the signal component has an effect on the correlation between black pattern and genetic distance. The significance of the correlations and partial correlations was calculated by comparing the actual $r$-statistic to a distribution of $r$-scores based on 1000 random permutations of the distance matrices.

## Results

### *Mitochondrial variation*

We found a total of 13 haplotypes across all regions and years. The total nucleotide diversity (Pi) was 0.00030, and the average nucleotide difference ($k$) was 0.19042. The number of haplotypes/region in each year ranged from one (Estonia 2011) to four in most regions. The average number of segregating sites per region in each year was 2.3, whereas the overall haplotype diversity (H) and nucleotide diversity (Pi) were 0.16940 and 0.00028, respectively. The haplotype network identified

one main haplogroup, which included samples from all regions, years and both morphs (Fig. 2). A second major haplotype was shared between samples of both morphs from Finland, and the most distinct haplotypes overall correspond to Estonian samples (Fig. 2).

The analysis of molecular variance (AMOVA) indicated that only 5.4% of the variation is explained by differences between years (Table S2, Supporting information), whereas 99.01% is explained within years. A negative variance component (−4.45) was found between morphs within years. Usually, negative variance components in an AMOVA indicate an absence of genetic structure because the true value of the parameter estimate is zero. Hence, if the expectation of the estimator is zero, slightly positive or slightly negative variance components can occur by chance (Weir & Cockerham 1984; Weir & Hill 2002). As most of the haplotypes were shared between the morphs and years, we performed the Tajima's D test of neutrality on both morphs separately and in combination within regions, the exemption being the samples from Åland for which no variation was found neither within nor between morphs. Hence, the D statistic could not be calculated. For the rest of the samples, the minimum D value was −0.013 corresponding to Estonia, whereas the maximum value of −0.992 was observed in the yellow morph from Finland. None of the values were statistically significant for the yellow or white morphs in any combination, which indicates that the COI fragment can be considered as neutrally evolving.

### *Microsatellite variation*

Significant deviations from Hardy–Weinberg equilibrium (HWE) were observed within regions after Bonferroni corrections. However, none of the deviations occurred consistently at the same loci in all regions. Likewise, null alleles were predicted for most of the loci segregating at variable frequencies within regions. Null allelic frequencies ranged from 0.0001 to 0.044 with a mean of 0.015 per locus and region. Given such relatively low null allelic frequencies and random HWE

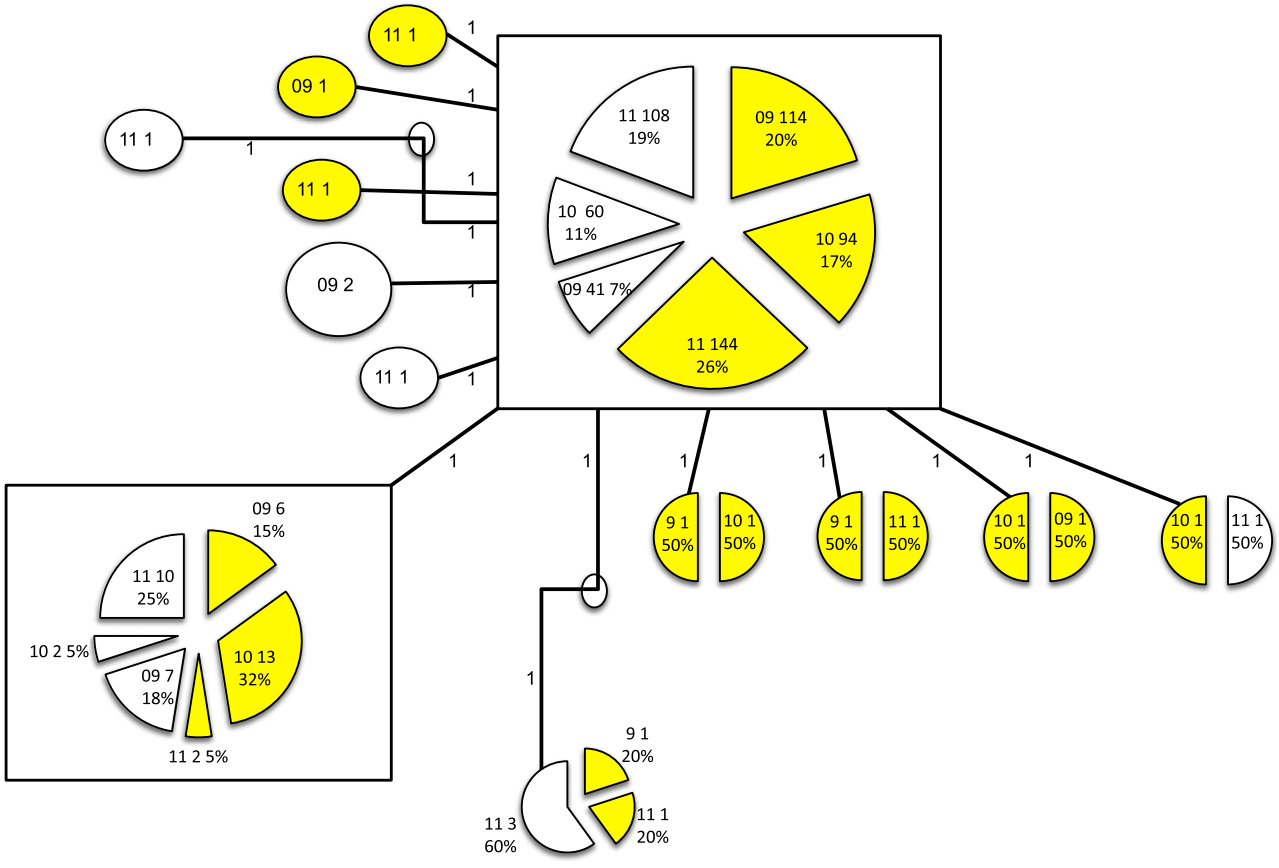


**Fig. 2** Mitochondrial COI haplotype network of *Parasemia plantaginis*. Squares indicate the main haplogroups. The colour in the slices represents colour morph. Numbers in the slices indicate the year, the number of samples within that year and its contribution (in percentage) to the haplotype. Numbers next to connecting lines indicate the number of mutational steps between the haplotypes. Empty circles represent undetected intermediate haplotypes.

deviations observed within loci and regions, all loci were kept for further analyses. No evidence of linkage disequilibrium was found between any locus pair.

The level of genetic diversity at microsatellite loci as estimated by the number of alleles per locus (Na) and the effective number of alleles (Ar), as well as by expected heterozygosity ($H_E$) is presented in Table 1. Within Åland, the white morph had significantly higher genetic variability than the yellow morph, whereas in Finland, the yellow morph showed equal or moderately higher variability across most variables measured. The same was true for the samples originating from the Alps. No between-morphs comparisons are available for Estonia because few yellow morphs (<3% of all samples) have been found there. However, between-year comparisons in Estonia showed relatively higher variability in 2009 (Table 1). All of the assignment tests failed to assign the totality of the samples back to their original morph category. The highest assignment success was observed in Åland where 84% of white morph samples could be assigned back to the white morph category (Table 1). No clear pattern was evident regarding private alleles in each morph, and low frequency alleles were more abundant every other year (Table 1).

The AMOVA results showed that most of the genetic variation within a given region and morph occurs within sampling sites (range: 79–98%; Table S3, Supporting information). No temporal differentiation was found in the Alpine region for neither morph. Moreover, no significant differentiation was observed between sampling sites within years (Table S3, Supporting information). The Bayesian clustering method implemented by STRUCTURE showed that the mean probability of the log-likelihood values (LK) saturated at $K = 19$ (Fig. S4, Supporting information). However, the *ad hoc* method of Evanno, which is based on the rate of change of the log-likelihood probabilities ($\Delta K$), indicated that $K = 2$. Although these results were consistent across all 10 iterations (Fig. S4, Supporting information), they must be taken with caution because STRUCTURE cannot accommodate temporal data. Hence, their biological interpretation is not straightforward. In our case, the $K = 2$ result shows no differentiation between the morphs within regions, whereas temporal differences exist for most of the regions (Fig. 3). For instance, both morphs from Åland 2009 clustered together, whereas a different cluster was observed in 2010–2011 in both morphs. This was the case also for the Estonian 2009 samples clustering separate from 2011. Less clear results, but within the same trend, were obtained from both Finnish morphs in 2009, showing different clustering from 2010 and 2011, which in turn constituted its own cluster. No different clustering was observed, however, for the Alpine morphs between 2009 and 2010. Congruently, none of the $F_{ST}$ comparisons showed statistical differences between morphs within years (Table 1). The same was true for the Jost's D and Nei's D unbiased estimators (Table S4, Supporting information). However, there were significant differences within morphs across years in all sampling regions (Table 1). The proportion of shared alleles was in line with $F_{ST}$ comparisons showing no differences within morphs for the same periods (Table 1). Lastly, as with the STRUCTURE analysis, the long-term connectivity results from MIGRATE should be taken with caution due to the temporal nature of our data. For instance, the migration rate parameter ($M$) shows cases of migration to past years (Table 2). In interpreting its biological meaning, this analysis suggests that regions have been either historically replenished locally, as in the case of the Alps, or exchanging migrants with their immediate neighbouring regions, as in the case of Finland with Åland and Estonia. Interestingly, migrate exchange between both morphs appeared not to be assortative according to coloration.

### *Warning signal variation*

In the forewing, 75.35% of the variance was explained by the first principal component. This component distinguished the morphs by high (positive) values of LW and MW, as well as by low (negative) values of UV and SW. The second principal component explained 18.88% of the total variance distinguishing between morphs by high MW and SW values, as well as by low UV and LW values (Fig. S2, Supporting information). In the hindwing, only high values of LW differentiated the morphs in the first principal component explaining 82.58% of the variance, whereas the second principal component explained only 8.6% of the variance (Fig. S2, Supporting information). As most of the variation could be explained by the first principal component in both wing sets, we used this component's scores as a synthetic variable representative of wing colour for further analyses.

The analysis of variance (ANOVA) showed mixed results within and between the morphs across the different regions and years. For the forewing, the UV varied between and within the morphs every other year (except in the Alps for which only two years were sampled) (Fig. 4). However, in the Alps, significant differences were observed between the morphs and years for the rest of the wavelengths measured. Likewise, lower values in the MW and LW across regions were observed in 2010, whereas the SW showed higher values for the same year (Fig. 4). For the hindwing, the UV-reflectance varied significantly between the morphs across regions and years, whites having higher UV-reflectance than

**Table 1** Microsatellite genetic diversity indices of *Parasemia plantaginis* morphs

| | N | | Na | | Ar | | He | | PA | | Assign | | PSA | PW $F_{STW}$ | $F_{SR}$ | | PW $F_{STB}$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Y | W | Y | W | Y | W | Y | W | Y | W | Y | W | YW | YW | Y | W | Y | W |
| Åland | | | | | | | | | | | | | | | | | | |
| 2009 | 35.7 | 10 | 9.7 | 10 | 5.1 | 5.55 | 0.80 | 0.86** | 6 | 0 | 0.78 | 0.30 | 0.271 | 0.003 | 0.069 (9,10,11) | 0.042 (9,10,11) | 0.148* (9−10) | 0.177* (9−10) |
| 2010 | 12.2 | 14 | 6.3 | 8** | 3.8 | 5.44** | 0.60 | 0.84** | 0 | 0 | 0.53 | 0.84 | 0.460 | 0.010 | | | −0.004 (10–11) | 0.047 (10–11) |
| 2011 | 5.5 | 5.0 | 4.6 | 6.0** | 3.5 | 4.1** | 0.70 | 0.84** | 0 | 0 | 0.16 | 0.40 | 0.289 | −0.023 | | | 0.138* (9–11) | 0.247* (9–11) |
| Finland | | | | | | | | | | | | | | | | | | |
| 2009 | 44 | 89 | 16.4 | 18.3** | 7.9** | 6.86 | 0.79 | 0.77 | 24 | 29 | 0.57 | 0.71 | 0.284 | 0.001 | 0.024 (9,10,11) | 0.094 (9,10,11) | 0.042* (9–10) | 0.042* (9–10) |
| 2010 | 42.5 | 19.1 | 9.3** | 7.9 | 4.9 | 4.8 | 0.66 | 0.66 | 1 | 0 | 0.65 | 0.52 | 0.307 | 0.001 | | | 0.007 (10–11) | −0.002 (10–11) |
| 2011 | 20.3 | 44.3 | 8.3 | 11** | 4.8 | 4.5 | 0.65 | 0.65 | 0 | 9 | 0.54 | 0.51 | 0.249 | 0.008 | | | 0.050* (9–11) | 0.050* (9–11) |
| Alps | | | | | | | | | | | | | | | | | | |
| 2009 | 28 | 14.7 | 12.5** | 9.4 | 7.3** | 6.5 | 0.72 | 0.70 | 16 | 0 | 0.44 | 0.26 | 0.264 | −0.011 | 0.001 (9,10) | 0.015 (9,10) | −0.001 (9–10) | −0.002 (9–10) |
| 2010 | 18.7 | 9.9 | 9.9** | 6.6 | 6.2** | 4.8 | 0.69 | 0.68 | 0 | 0 | 0.52 | 0 | 0.399 | −0.021 | | | | |
| Estonia | | | | | | | | | | | | | | | | | | |
| 2009 | 23.3 | | 10.6 | | 6.7 | | 0.774** (9–11) | | 1 | | 0.91 | | 0.272 (9–11) | | | 0.066 (9,10,11) | 0.042* (9–11) | |
| 2011 | 33.3 | | 9.2 | | 9.2** (9–11) | | 0.662 | | 13 | | 0.91 | | | | | | | |

Y, Yellow morph; W, white morph; N, average number of individuals genotyped per locus; Na, average number of alleles per locus; Ar, average number of effective alleles per locus; He, average expected heterozygosity; PA, private alleles; Assign, proportion of individuals assigned to its own population; PW $F_{STW}$, pairwise $F_{ST}$ within years; $F_{SR}$, allelic differentiation between sampling sites within years (via AMOVA); PW $F_{STB}$, Pairwise $F_{ST}$ between years; PSA, proportion of shared alleles.

*$P < 0.05$. Numbers in parentheses indicate the years compared.

**Significant after *t*-test.

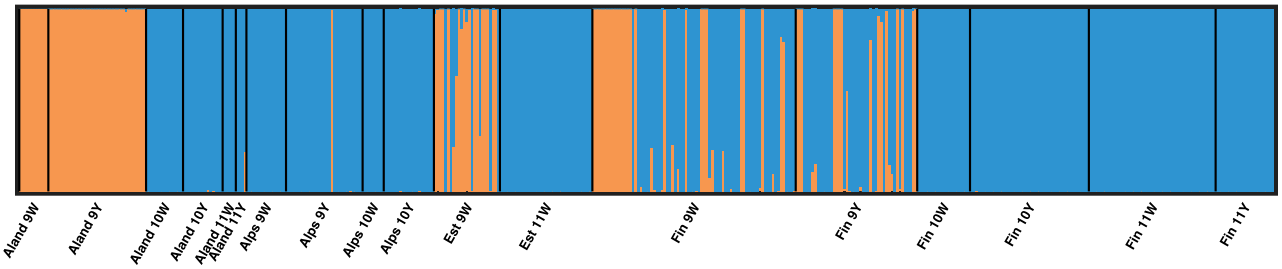


**Fig. 3** Estimated probabilities of population membership assessed by multilocus microsatellite genotypes of *Parasemia plantaginis*. Each sample is represented by a vertical bar. Colours represent the population been assigned. The name of the population indicates the region, the year of sampling and the colour morph.

**Table 2** Migration rate (*M*) observed between regions in terms of effective migrants per generation (4 *Nm*)

| *M* | 2.5% | 25.0% | Mode | 75% | 97.5% | Median | Mean |
|---|---|---|---|---|---|---|---|
| Fin09W->Aland09W | 10.00 | 22.00 | 30.33 | 38.00 | 49.33 | 31.00 | 45.2 |
| Fin10Y->Aland09Y | 4.67 | 15.33 | 23.67 | 30.67 | 41.33 | 24.33 | 35.41 |
| Fin09Y->Aland10W | 0.00 | 10.00 | 17.67 | 26.00 | 36.00 | 19.67 | 26.88 |
| Fin09W->Aland10Y | 26.00 | 38.67 | 47.00 | 54.67 | 66.67 | 47.00 | 69.34 |
| Fin11Y->Aland11W | 0.00 | 6.00 | 13.67 | 20.00 | 30.67 | 15.67 | 20.43 |
| Fin10W->Aland11Y | 0.00 | 8.67 | 16.33 | 23.33 | 32.00 | 17.00 | 24.05 |
| Alps10Y->Alps09W | 23.33 | 34.00 | 42.33 | 50.00 | 62.00 | 43.67 | 63.95 |
| Alps10W->Alps09Y | 8.67 | 21.33 | 30.33 | 38.00 | 50.67 | 31.00 | 45.01 |
| Alps09Y->Alps10W | 8.00 | 19.33 | 27.67 | 34.67 | 46.00 | 28.33 | 41.36 |
| Alps09Y->Alps10Y | 8.67 | 19.33 | 27.00 | 34.67 | 45.33 | 27.67 | 40.95 |
| Fin10Y->Est09W | 2.67 | 14.00 | 22.33 | 30.00 | 41.33 | 23.00 | 33.59 |
| Fin09W->Est11W | 14.67 | 28.00 | 37.00 | 45.33 | 58.67 | 37.67 | 55.3 |
| Fin09Y->Fin09W | 16.67 | 27.33 | 35.00 | 42.67 | 52.67 | 35.67 | 52.75 |
| Fin09W->Fin09Y | 34.67 | 47.33 | 56.33 | 64.00 | 76.67 | 57.00 | 84.17 |
| Fin11Y->Fin10W | 10.67 | 22.00 | 30.33 | 37.33 | 48.67 | 31.00 | 45.21 |
| Fin09W->Fin10Y | 16.00 | 28.67 | 37.67 | 45.33 | 58.00 | 38.33 | 56.36 |
| Fin09W->Fin11W | 19.33 | 32.00 | 40.33 | 48.67 | 61.33 | 41.00 | 60.84 |
| Fin09W->Fin11Y | 15.33 | 28.67 | 37.00 | 45.33 | 58.00 | 37.67 | 55.67 |

The multilocus likelihood of the percentiles for *M* is shown. Migration from the donor to the receiver population is indicated by->. The region, year and the colour morph are indicated in the last three digits of the population name.

yellows in the Alps and Finland, but no temporal differences were observed within the morphs. The exception was in Åland where the UV-reflectance showed no temporal differentiation within or between the morphs (Fig. 5). The Alps and Finland, on the other hand, showed significant differences between morphs and years for most of the wavelengths measured. Interestingly, the white morph showed no variation across years, whereas the yellow morph showed temporal variability. In both wing sets, no temporal differences existed in the signal size within or between the morphs, except in Finland 2011 where the morphs differed significantly in their hindwing signal size (data not shown). The white morph had less melanization in the hindwings compared with the yellow morph. Finally, the MW showed the least temporal variation in all regions within and between morphs for both wing sets.

### *Combined genetic and phenotypic results*

Overall, for both wing sets, we found low or negative correlation values in the Mantel and partial Mantel tests performed for the nuclear and mitochondrial data, indicating that most of the signal components do not evolve in a neutral fashion (Table 3). Significant positive correlations, however, were observed between both neutral markers and the forewing's signal size in the yellow morph from the Alps 2010. The correlations remained significant when the effect of the pattern was controlled in the partial correlations. Congruently, the hindwing also showed a significant positive correlation between the signal size and the microsatellite data set for the same sample, as did the signal component Col (Table 3). This indicates that the size of the warning signal, and possibly the coloration, could be evolving in a

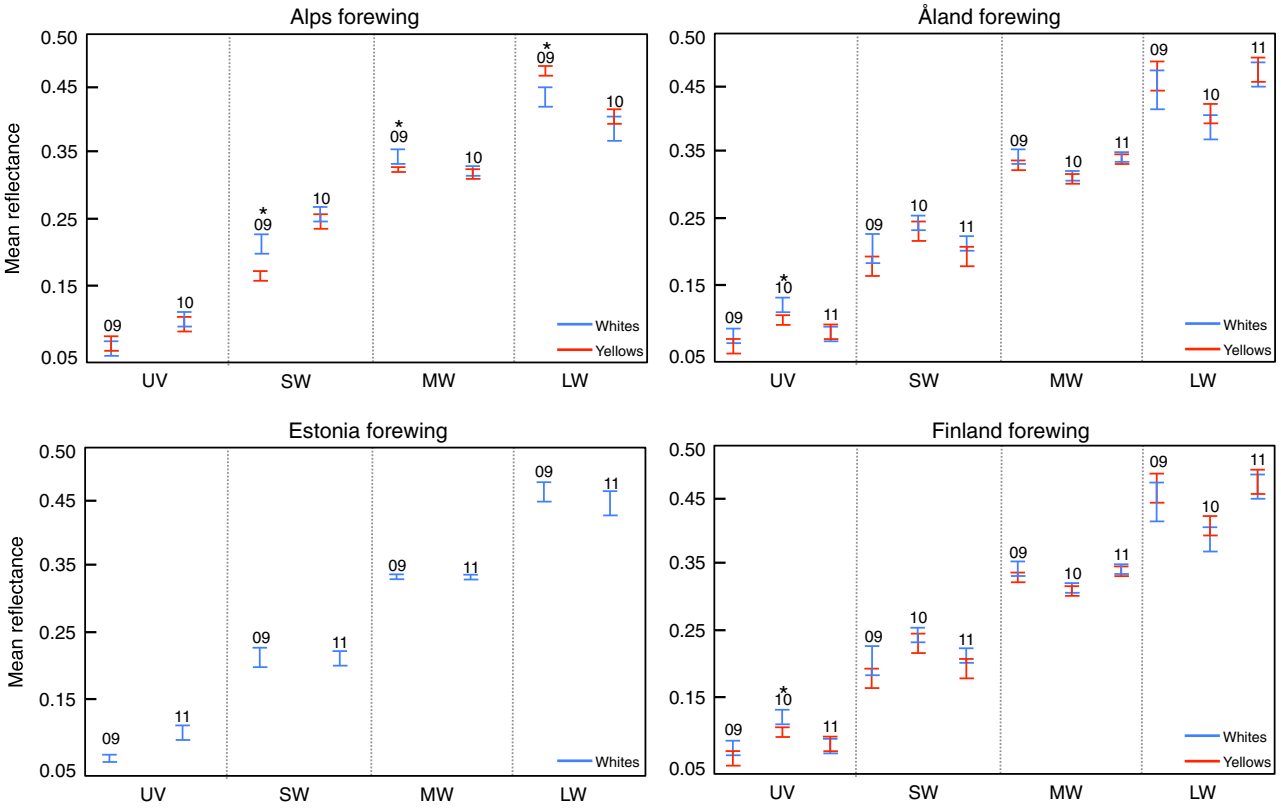


**Fig. 4** Comparison of forewing saturation values of ultraviolet (UV approximately 300–400 nm), short (SW approximately 400–500 nm), medium (MW approximately 500–600 nm) and long wavelengths (LW approximately 600–700 nm) within and between *Parasemia plantaginis* colour morphs in each of the sampling regions. Numbers indicate the year being compared. Asterisks indicate significant difference ($P < 0.05$).

neutral fashion in the Alpine yellow morphs in 2010. Interestingly, a significant negative correlation was also observed in the yellow morph between the forewing pattern and the COI marker in the Alps 2009, and it remained negative after controlling for the effect of the signal in the partial mantel tests (Table 3). This indicates a clear non-neutrality of the forewing pattern. Finally, significant negative correlations were also present in the partial mantel tests between the COI marker, the signal size and UV-reflectance while controlling for the forewing pattern in the white morph from Estonia and Finland 2011 (Table 3). This suggests that the forewing pattern does not have a significant effect in the non-neutral evolution of these signal components.

## Discussion

Here, we investigated the spatio-temporal relationship between genetic and warning signal variation of two sympatric colour morphs across four regions. Our results showed that the two morphs constitute a single genetic population within a given region and year. However, the genetic structure of the population changes temporally, and temporal variation in coloration and UV-reflectance occurs differently for the hind- and forewings of both morphs. Other components such as the size of the warning signal size do not vary temporally between wing sets or morphs. We also found indications of non-neutral evolution for all warning signal components analysed except for the signal size. Our results suggest that the elements that constitute a warning signal can experience different selection pressures within a single genetic population. Temporal fluctuations in selection pressures and genetic structure could be a likely mechanism for the maintenance of local colour polymorphisms.

### *Temporal genetic variation*

Overall, we found a relatively high genetic variability in the nuclear microsatellite loci, which is in line with previous studies of other moth species (Simonato *et al.* 2007; Magalhaes *et al.* 2011; Voudouris *et al.* 2012). On the other hand, the observed genetic variability in the COI was low to moderate in comparison with other moth studies in Europe for the same mitochondrial

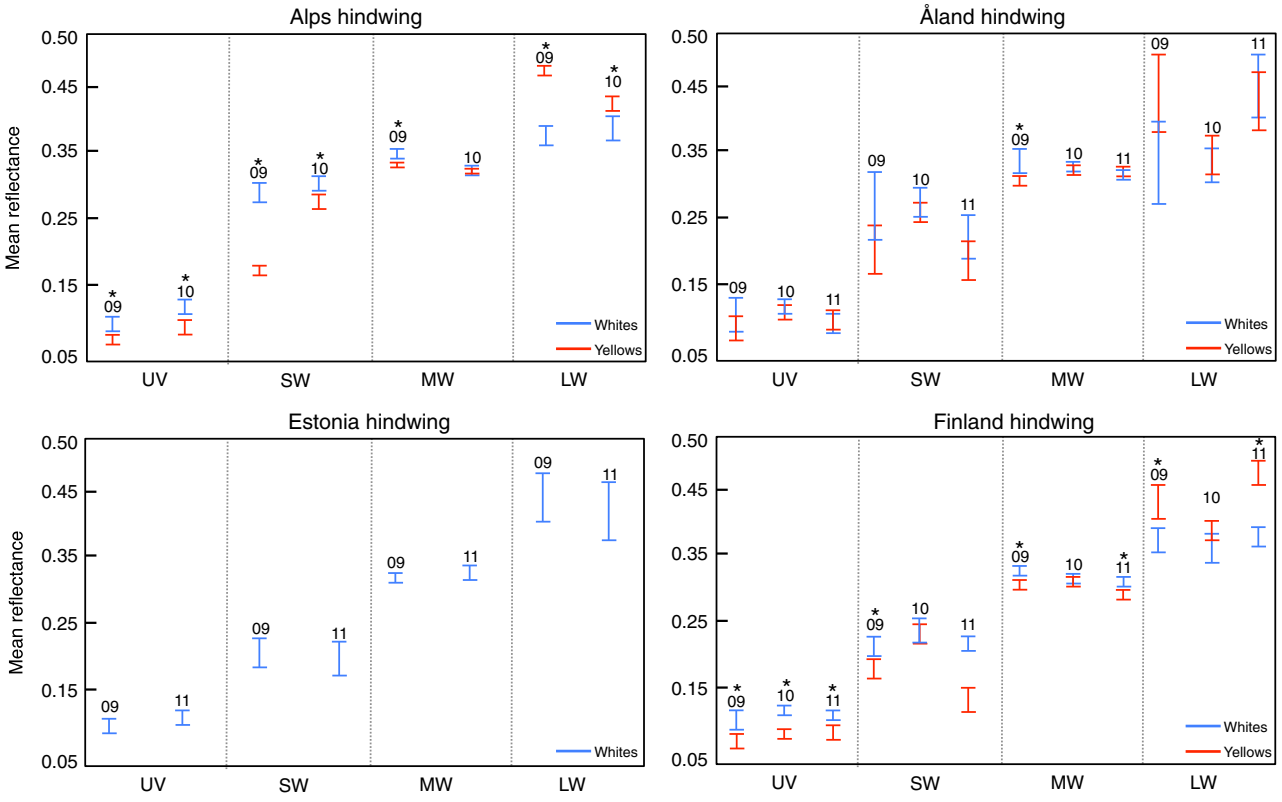


**Fig. 5** Comparison of hindwing saturation values of ultraviolet (UV approximately 300–400 nm), short (SW approximately 400–500 nm), medium (MW approximately 500–600 nm) and long wavelengths (LW approximately 600–700 nm) within and between *Parasemia plantaginis* colour morphs in each of the sampling regions. Numbers indicate the year being compared. Asterisks indicate significant difference ($P < 0.05$).

region (Meraner *et al.* 2008; Thaler *et al.* 2008; Kerdelhué *et al.* 2009; Rousselet *et al.* 2011). Nonetheless, given *P. plantaginis*' wide geographical distribution, the relatively low haplotypic variability does not seem to be a limiting factor for the successful persistence of the species. This has shown possible for other species with low or nil mitochondrial variation such as lemmings, elephant seals and termites, whose ecological success does not seem to be impaired by the lack of variation in their mitochondrial genomes (Broughton & Grace 1994; Fedorov & Stenseth 2001; Hoelzel *et al.* 2002).

Our results for the COI data set showed no clear temporal fluctuation in haplotype diversity as two main haplotypes were shared by both morphs across years and regions (Fig. 2). However, the microsatellite data set showed temporal differences in allelic frequencies in both morphs (Table 1). This discrepancy between markers could be due to ecological or evolutionary processes. For instance, dispersal differences exist between males and females, with males having better flying capabilities (authors' Personal observation). Sex-biased dispersal is likely to generate discordant patterns between nuclear and mitochondrial markers, as it has been shown in several other species with high-vagility males and females with restricted dispersal (Castella *et al.* 2001; Toews & Brelsford 2012). Alternatively, due to the non-recombining nature and unipaternal inheritance of haploid mtDNA, lineage sorting is completed faster than nuclear, reflecting more ancient variation (Funk & Omland 2003). Thus, our results could suggest episodes of greater historic connectivity, which have not been erased by current dispersal or landscape alterations.

Apart from the long-debated (see Jones 1989; Skipper 2009 for reviews) and largely unresolved (but see Goulson 1999; O'Hara 2005) drift vs. equilibrium case of the Scarlet tiger moth (*Callimorpha dominula)* from Cothill, scarce information is available about temporal genetic variation in moths, and in Lepidopterans in general. In one example, an allozyme-based study examined allelic frequency variation in the ermine moth (*Yponomeuta padellus*) over a 13-year period (Raijmann & Menken 2000). Significant differences were found within the short term (four years), but not in the long term, where allelic frequencies remained relatively stable. Similarly, populations of the violet copper butterfly (*Lycaena helle*) were compared after a 15-year period using microsatellites

**Table 3** Mantel tests indicating the correlation between neutral genetic patterns (Gen) and those of the different signal (Sig) components; colour (Col), signal size (Ss) and ultraviolet reflectance (UV)

| Region | Gen vs. Sign | Gen vs. Sign vs. Patt | Gen vs. Patt | Gen vs. Patt vs. Sign |
|---|---|---|---|---|
| A | | | | |
| Alps 09 Y Col | 0.1482 | 0.1715 | −0.1714* | −0.1917 |
| Alps 09 Y Ss | −0.0408 | −0.0277 | −0.1714* | −0.1688 |
| Alps 09 Y UV | 0.1274 | 0.1035 | −0.1714* | −0.1547 |
| Alps 10 Y UV | −0.0702 | −0.0651 | 0.1288 | 0.1241 |
| Alps 10 Y Col | 0.0015 | −0.034 | 0.1288 | 0.1321 |
| Alps 10 Y Ss | 0.5575* | 0.5477* | 0.1288 | 0.0250 |
| Est 11 W Col | 0.1325 | 0.1324 | −0.0313 | −0.0221 |
| Est 11 W Ss | −0.1380 | −0.1321* | −0.0313 | −0.0315 |
| Est 11 W UV | −0.2450 | −0.2449* | −0.0313 | 0.0074 |
| Fin 11 W Col | −0.0107 | −0.0031 | 0.1034 | 0.1029 |
| Fin 11 W Ss | −0.0937 | −0.0844* | 0.1034 | 0.0950 |
| Fin 11 W UV | 0.0762 | 0.0792 | 0.1034 | 0.1056 |
| B | | | | |
| Alps 09 Y Col | −0.0258 | −0.0209 | −0.1377 | −0.1369 |
| Alps 09 Y Ss | −0.0860 | −0.0760 | −0.1377 | −0.1318 |
| Alps 09 Y UV | −0.0381 | −0.0608 | −0.1377 | −0.1455 |
| Alps 10 Y UV | −0.0549 | −0.0494 | 0.1332 | 0.1310 |
| Alps 10 Y Col | −0.0495 | −0.0333 | 0.1332 | 0.1281 |
| Alps 10 Y Ss | 0.2319* | 0.2122* | 0.1332 | 0.0931 |
| Est 11 W Col | 0.0140 | 0.0147 | 0.0617 | 0.0618 |
| Est 11 W Ss | 0.0569 | 0.0511 | 0.0617 | 0.0564 |
| Est 11 W UV | −0.0084 | −0.0031 | 0.0617 | 0.0612 |
| Fin 11 W Col | −0.0589 | −0.0586 | −0.0494 | −0.0490 |
| Fin 11 W Ss | 0.0067 | 0.0018 | −0.0494 | −0.0490 |
| Fin 11 W UV | 0.0848 | 0.0837 | −0.0494 | −0.0475 |
| C | | | | |
| Alps 09 Y Col | −0.0357 | | | |
| Alps 09 Y Ss | 0.0739 | | | |
| Alps 09 Y UV | 0.0341 | | | |
| Alps 10 Y UV | −0.1260 | | | |
| Alps 10 Y Col | 0.0346 | | | |
| Alps 10 Y Ss | −0.0544 | | | |
| Est 11 W Col | 0.0527 | | | |
| Est 11 W Ss | −0.0521 | | | |
| Est 11 W UV | 0.1412 | | | |
| Fin 11 W Col | 0.1051 | | | |
| Fin 11 W Ss | 0.0194 | | | |
| Fin 11 W UV | 0.0289 | | | |
| D | | | | |
| Alps 09 Y Col | 0.0332 | | | |
| Alps 09 Y Ss | −0.0107 | | | |
| Alps 09 Y UV | −0.0244 | | | |
| Alps 10 Y UV | 0.0254 | | | |
| Alps 10 Y Col | 0.2104* | | | |
| Alps 10 Y Ss | 0.1496* | | | |
| Est 11 W Col | 0.0508 | | | |
| Est 11 W Ss | 0.0164 | | | |
| Est 11 W UV | 0.0164 | | | |
| Fin 11 W Col | 0.0486 | | | |
| Fin 11 W Ss | −0.0412 | | | |
| Fin 11 W UV | 0.0564 | | | |

Panels A and B show results from COI gene fragment and microsatellite loci, respectively, in the forewing. Also shown are the partial Mantel correlations between genetic and signal components while controlling for the effect of pattern (Patt). Panels C and D show results from COI gene fragment and microsatellite loci, respectively, in the hindwing. Names indicate region, year, colour morph and signal component.
*$P \leq 0.05$.

(Habel *et al.* 2011). The results revealed that site-specific alleles were identical in most cases after the 15-year period. In the present study, our microsatellite data indicated temporal variation for both morphs. The genetic structure changed every second year within morphs and regions as evidenced by significant $F_{ST}$ *b* values (Table 1). Accordingly, with the exception of Åland, new alleles were present every second year in all regions (Table 1). It is unpredictable, however, how allelic frequencies would compare to previous findings if sampling had been performed over a longer time period (i.e. >10 years). Nevertheless, it could be expected that allelic frequencies follow cyclic fluctuations induced by continuing ecological or climatic changes occurring within regions. For instance, recent work suggests that changes in predator community composition and alternate prey abundance can shift local selection regimes in *P. plantaginis* (Nokelainen *et al.* 2014). Likewise, seasonality significantly alters gene frequencies in the wild due to phenological shifts in the pine processionary moth (*Thaumetopoea pityocampa*) (Santos *et al.* 2011). Thus, long-term equilibrium of allelic frequencies might be approached through stabilizing selection in spite of short-term fluctuations in *P. plantaginis*.

Temporal genetic fluctuations may also be observed if (i) populations follow a metapopulation dynamics, (ii) biased mating preferences occur within the population and (iii) through effective immigration of individuals with different genetic background (Austin *et al.* 2011; Hanski *et al.* 2011). We have observed the complete absence of *P. plantaginis* in locations that have had high abundances in the previous year. Likewise, neighbouring empty locations often show a sudden presence of the species, resembling the source-sink dynamics of a metapopulation. However, more field data are needed to confirm this hypothesis, especially on the temporal genetic structure of the more sedentary and elusive females. On the other hand, mating experiments have shown no clear fitness benefits for the offspring when females were mated with both colour morphs (F. Santostefano, J. A. Galarza & J. Mappes, in preparation). This suggests that other cryptic cues (i.e. pheromone variation) could have a more relevant role than coloration in mating preferences within the population and warrants further investigation. Given our data, it is more likely that temporal genetic structure is the result of immigration from neighbouring populations within the region. If the individuals analysed within each sampling site had been residents (i.e. hatched from overwintering larvae), we would expect the genetic composition of the population to be more homogeneous from year to year. Future sampling of males in the beginning and end of the flying season should corroborate this.

### *Temporal variation in warning signal components*

It is widely thought that temporal variation in selection influences the maintenance of variation in warning signals (Endler & Mappes 2004; Mappes *et al.* 2005; Valkonen *et al.* 2011). Our temporal analyses of warning signal variation showed contrasting results between the fore- and hindwings for most of the components. In the hindwing, the UV-reflectance showed no temporal differences within the morphs in any of the regions, whereas it varied temporally in the forewing of both morphs (Figs 4 and 5). This suggests that UV-reflectance in the hindwing could be near optima (i.e. fixed), probably due to strong selection. It is not known, however, whether UV-reflectance in *P. plantaginis* functions as a warning signal or whether it has another function such as intraspecific signalling. In such case, different selection mechanisms could lead to the observed lack of variation. Results from other Lepidoptera studies, which have manipulated wing UV-reflectance of artificial prey, suggest that UV-reflectance is not an effective warning signal, but rather attractive to avian predators (Lyytinen *et al.* 2001, 2004). Conversely, more recent findings showed that UV-reflectance plays a significant role in mate choice in Lepidoptera (Kemp 2008).

Changes in coloration (SW, MW, LW) followed the same trend as the UV-reflectance, fluctuating more in the forewing of the yellow morph and less in the hindwing of the white morph (Figs 4 and 5). Such temporal changes in coloration have no clear relation with the genetic fluctuations observed. The coloration fluctuates irrespectively of whether the genetic composition of the population has changed or not. This could indicate that coloration is highly sensitive to local ecological processes, particularly those experienced during development. It has been shown that the warning signals of *P. plantaginis* larvae (an orange patch against a dark body) are influenced by temperature and diet composition (Lindstedt *et al.* 2009). Moreover, other studies have found that microclimatic developing conditions can significantly alter moth adult phenotype frequencies in the wild (Goulson & Owen 1997; Jones 2000). The lower temporal fluctuations observed in the hindwing could be expected as coloration is the hindwing's primary aspect involved in signalling against predators (Nokelainen *et al.* 2012; Hegna *et al.* 2013). Hence, it could be less subject to variability due to strong directional selection. Interestingly, more hindwing colour fluctuations occurred in the yellow morph than in the white morph (Figs 4 and 5). This could be the outcome of different sources of selection acting. In *P. plantaginis*, females seem to prefer white males, but yellow males appear to have

a more efficient warning signal against predators (Nokelainen *et al.* 2012). Thus, sexual selection could play a bigger role in the fitness of the white morph, whereas directional selection could have greater fitness effects in the yellow morph. Alternatively, it is likely that differences in pigmentation costs exist between white and yellow coloration. It is also plausible that within-morph variation in coloration is not very important to predators, in particularly for generalist predators (Ihalainen *et al.* 2012), which tend to generalize their behaviour towards warning signals that are similar enough (Ham *et al.* 2006). This may allow for small variation in the warning signal without hindering signal efficacy.

In moths, forewings are always visible when they rest, whereas the hindwings are often hidden by the forewings. Therefore, forewings tend to be cryptic, whereas hind wings often have evolved warning coloration that moths suddenly unleash to frighten their predators. In some Arctiidae, however, forewings have been suggested to have disruptive pattern elements (e.g. in genus *Grammia*) (Simmons 2009). *Parasemia plantaginis* forewing patterning is fairly similar to *Grammia*'s forewings, and it is reasonable to ask whether they have a protective function, aposematic or disruptive. In this regard, recent experiments by Hegna & Mappes (2014) did not find evidence of disruptive function of forewings when moth dummies were presented on green background to predators. Instead, their results suggested that the forewing patterning is aposematic. However, A. Honma, J. K. Valkonen & J. Mappes (in preparation) found some support of a disruptive function based on detection delays of bird predators when searching *P. plantaginis* in backgrounds that resembled litter of meadow. This is congruent with our filed observations of *P. plantaginis* faking dead and dropping down when approached. Therefore, we suggest that the protective function of forewing pattern can be dual: warning and disruptive depending on the background.

Only one significant difference was found between the morphs' hindwing signal size in the Finnish region in 2011. This indicates that stabilizing selection could occasionally affect the morphs differently. Likewise, the signal size showed a non-neutral pattern in both wing sets of the Alpine yellow morphs in 2010, but not in the previous year, where it appeared as neutrally evolving (Table 3). This apparent temporal shift in selection is interesting because melanization has direct effects in thermoregulation in the larval and adult stages constraining signal size, and as a consequence, its efficacy (Lindstedt *et al.* 2009; Hegna *et al.* 2013). Hence, alternating selection between thermoregulation and signal size could partly explain local differences in the evolution of signal size.

## Conclusion

In spite of the relatively short duration of this study, the analyses in different regions and repeated sampling of the same populations within these regions provide evidence that (i) some warning signal components vary through time, whereas others are fixed, (ii) temporal variation in signal components occurs differently in the fore- and hindwings, (iii) random mating occurs between the morphs, (iv) local genetic structure fluctuates and (v) warning signal variation cannot be related to genetic temporal variation. Our results evidence the complex interplay between local factors (i.e. predation, gene flow and the environment) and their role in counteracting the homogenizing effect of directional selection on aposematic species.

In sum, there is no simple answer to how the colour polymorphism is maintained in *P. plantaginis*. Based on this and our previous results, we believe there are two main factors operating in parallel. On the one hand, the trade-off between sexual selection and survival selection giving an advantage for whites in mating but favouring yellows against bird predators (Nokelainen *et al.* 2012) is certainly an important element but hardly sufficient to explaining the long-term coexistence of both morphs. On the other hand, we have found that selection by predators promotes the development of a small-scale selection mosaic (Nokelainen *et al.* 2014). The efficacy of the warning signal depends on predator community composition where some species (e.g. Paridae) avoid warning signals, whereas some others do not (e.g. Prunellidae). In addition, as chemical defences are more costly for yellows (Nokelainen *et al.* 2012), whites are expected to obtain an advantage when naïve predators dominate in the predator community. In high latitudes, the predator community composition is likely to change yearly and is dominated by migratory passerines, which impedes the evolution of an optimal local defence strategy. This suggests that *P. plantaginis* is fundamentally different to other well-known aposematic species such as the *Heliconius* system, which is a classic example of aposematism and mimicry (Ruxton *et al.* 2004; Mappes *et al.* 2005) where predators favour local morphs selecting against new colour variants and dispersal (Mallet & Barton 1989; Langham 2004). Finally, in the present study, we found large panmictic populations with yearly fluctuations in their genetic composition. This indicates that high gene flow between the morphs may counterbalance the different temporal selection pressures experienced by each morph, and hence allowing for coexistence at variable frequencies. Future studies should consider geographical differences in expression patterns of functional genes affecting warning signal components.

## Acknowledgements

Several assistants gave valuable contribution to field sampling: Tõnis Tasane, Madis Tasane, Katja Ojala, Tom Hoogesteger, Jonathan Hegna and Riikka Lämsä. Andrés Lopez-Sepulcre and Janne Valkonen provided valuable comments on the statistics. We also thank four anonymous reviewers for their constructive comments, and subject editor Prof. Rosemary Gillespie. Kaisa Suisto reared females for pheromone traps. Sari Viinikainen helped in laboratory work. Members of Darwin Club in the University of Jyväskylä, Bibiana Rojas, Carita Lindstedt, Swanne Gordon and Emily Burdfield-Steel, gave valuable comments on the manuscript. This study was funded by the Academy of Finland via the Centre of Excellence in Biological Interactions (Project: 21000002561), and the Biological Interactions Graduate School (RHH).

## References

Austin A, Ovaskainen O, Hanski I (2011) Size and genetic composition of the colonizing propagules in a butterfly metapopulation. *Oikos*, **120**, 1357–1365.

Beatty CD, Beirinckx K, Sherratt TN (2004) The evolution of Müllerian mimicry in multispecies communities. *Nature*, **431**, 63–66.

Beerli P (2006) Comparison of Bayesian and maximumlikelihood inference of population genetic parameters. *Bioinformatics*, **22**, 341–345.

Belkhir K, Borsa P, Goudet J, Chikhi L, Bonhomme F (1997) *Genetix v.4.01 logiciel sous Windows TM pour la génétique des populations*. Laboratoire Génome, Populations, Interactions CNRS UPR 9060, Montpellier, France.

Bohonak AJ (2002) IBD (isolation by distance): a program for analyses of isolation by distance. *Journal of Heredity*, **93**, 153–154.

Bowcock AM, Ruíz-Linares A, Tomfohrde J *et al.* (1994) High resolution human evolutionary trees with polymorphic microsatellites. *Nature*, **368**, 455–457.

Broughton RE, Grace JK (1994) Lack of mitochondrial DNA variation in an introduced population of the Formosan subterranean termite (Isoptera: Rhinotermitidae). *Sociobiology*, **24**, 121–126.

Brunton CF, Majerus ME (1995) Ultraviolet colours in butterflies: intra-or inter-specific communication? *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **260**, 199–204.

Castella V, Ruedi M, Excoffier L (2001) Contrasted patterns of mitochondrial and nuclear structure among nursery colonies of the bat *Myotis myotis*. *Journal of Evolutionary Biology*, **14**, 708–720.

Chapuis M-P, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, **24**, 621–631.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Cott HB (1940) *Adaptive Coloration in Animals*. Methuen & Co. Ltd, London, UK.

Davis AK, Chi J, Bradley C, Altizer S (2012) The redder the better: wing color predicts flight performance in monarch butterflies. *PLoS ONE*, **7**, e41323.

Dieringer D, Schlötterer C (2003) Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes*, **3**, 167–169.

Dubatolov V, de Vos R (2010) Tiger-moths of Eurasia (Lepidoptera, Arctiidae). *Neue Entomologische Nachrichten*, **65**, 1–106.

Earl D, vonHoldt B (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Edmunds M (1974) *Defence in Animals*. Longman Group Ltd, Harlow, UK.

Endler JA, Mappes J (2004) Predator mixes and the conspicuousness of aposematic signals. *American Naturalist*, **163**, 532–547.

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, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Fedorov VB, Stenseth NC (2001) Glacial survival of the Norwegian lemming (Lemmus lemmus) in Scandinavia: inference from mitochondrial DNA variation. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **268**, 809–814.

Funk DJ, Omland KE (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology Evolution and Systematics*, **34**, 397–423.

Galarza JA, Viinikainen SM, Ashrafi R, Mappes J (2011) First microsatellite panel for the Wood Tiger Moth (*Parasemia plantaginis*). *Conservation Genetics Resources*, **3**, 197–199.

Goudet J (1995) F_STAT (2.9.3) a program for IBM compatible PCs to calculate Weir and Cockerman's (1984) estimators of F-statistics. *Journal of Heredity*, **86**, 485–486.

Goulson D (1999) Why studies of the medionigra polymorphism in *Panaxia dominula* do not provide evidence for natural selection. *Oikos*, **87**, 181–182.

Goulson D, Owen DF (1997) Long-term studies of the medionigra polymorphism in the moth *Panaxia dominula*: a critique. *Oikos*, **80**, 613–617.

Grant PR, Grant BR (2002) Unpredictable evolution in a 30-year study of Darwin's finches. *Science*, **296**, 707–711.

Grill CP, Moore AJ (1998) Effects of a larval antipredator response and larval diet on adult phenotype in an aposematic ladybird beetle. *Oecologia*, **114**, 274–282.

Habel JC, Finger A, Schmitt T, Neve G (2011) Survival of the endangered butterfly *Lycaena helle* in a fragmented environment: genetic analyses over 15 years. *Journal of Zoological Systematics and Evolutionary Research*, **49**, 25–31.

Ham AD, Ihalainen E, Lindström L, Mappes J (2006) Does colour matter? The importance of colour in avoidance learning, memorability and generalisation. *Behavioral Ecology and Sociobiology*, **60**, 482–491.

Hanski I, Mononen T, Ovaskainen O (2011) Eco-evolutionary metapopulation dynamics and the spatial scale of adaptation. *The American Naturalist*, **177**, 29–43.

Hegna RH, Mappes J (2014) Influences of geographic differentiation in the forewing warning signal of the wood tiger moth in Alaska. *Evolutionary Ecology*, DOI: 10.1007/s10682-014-9734-7.

Hegna RH, Nokelainen O, Hegna JR, Mappes J (2013) To quiver or to shiver: increased melanization benefits thermoregulation, but reduces warning signal efficacy in the wood tiger moth. *Proceedings of the Royal Society B: Biological Sciences*, **280**, 20122812.

Hendry AP, Day T (2005) Population structure attributable to reproductive time: isolation by time and adaptation by time. *Molecular Ecology*, **14**, 901–916.

Hoelzel AR, Fleischer RC, Campagna C, Le Boeuf BJ, Alvord G (2002) Impact of a population bottleneck on symmetry and genetic diversity in the northern elephant seal. *Journal of Evolutionary Biology*, **15**, 567–575.

Ihalainen E, Rowland HM, Speed MP, Ruxton GD, Mappes J (2012) Prey community structure affects how predators select for Müllerian mimicry. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 2099–2105.

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.

Jones DA (1989) 50 years of studying the scarlet tiger moth. *Trends in Ecology & Evolution*, **4**, 298–301.

Jones DA (2000) Temperatures in the Cothill habitat of *Panaxia (Callimorpha) dominula* L. (the scarlet tiger moth). *Heredity*, **84**, 578–586.

Joron M, Mallet JLB (1998) Diversity in mimicry: paradox or paradigm? *Trends in Ecology & Evolution*, **13**, 461–466.

Jost L (2008) GST and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015–4026.

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: *Mamalian Protein Metabolism* (ed. Munro HN), pp. 21–132. Academic Press, New York.

Kemp DJ (2008) Female mating biases for bright ultraviolet iridescence in the butterfly Eurema hecabe (Pieridae). *Behavioral Ecology*, **19**, 1–8.

Kerdelhué C, Zane L, Simonato M *et al.* (2009) Quaternary history and contemporary patterns in a currently expanding species. *BMC Evolutionary Biology*, **9**, 220.

Kronforst MR, Young LG, Kapan DD *et al.* (2006) Linkage of butterfly mate preference and wing color preference cue at the genomic location of wingless. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 6575–6580.

Langham GM (2004) Specialized avian predators repeatedly attack novel color morphs of *Heliconius* butterflies. *Evolution*, **58**, 2783–2787.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA poly-morphism data. *Bioinformatics*, **11**, 1451–1452.

Lindstedt C, Lindström L, Mappes J (2009) Thermoregulation constrains effective warning signal expression. *Evolution*, **63**, 469–478.

Lindstedt C, Morehouse N, Pakkanen H *et al.* (2010a) Characterizing the pigment composition of a variable warning signal of *Parasemia plantaginis* larvae. *Functional Ecology*, **24**, 759–766.

Lindstedt C, Talsma JHR, Ihalainen E, Lindstrom L, Mappes J (2010b) Diet quality affects warning coloration indirectly: excretion costs in a generalist herbivore. *Evolution*, **64**, 68–78.

Lyytinen A, Alatalo RV, Lindström L, Mappes J (2001) Can ultraviolet cues function as aposematic signals? *Behavioral Ecology*, **12**, 65–70.

Lyytinen A, Lindström L, Mappes J (2004) Ultraviolet reflection and predation risk in diurnal and nocturnal Lepidoptera. *Behavioral Ecology*, **15**, 982–987.

Maan ME, Cummings ME (2009) Sexual dimorphism and directional sexual selection on aposematic signals in a poison frog. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 19072–19077.

Magalhaes IS, Gleiser G, Labouche A-M, Bernasconi G (2011) Comparative population genetic structure in a plant–pollinator/seed predator system. *Molecular Ecology*, **20**, 4618–4630.

Mallet J (2010) Shift happens! Shifting balance and the evolution of diversity in warning colour and mimicry. *Ecological Entomology*, **35**, 90–104.

Mallet J, Barton NH (1989) Strong natural selection in a Warning-Color hybrid zone. *Evolution*, **43**, 421–431.

Mappes J, Marples N, Endler JA (2005) The complex business of survival by aposematism. *Trends in Ecology & Evolution*, **20**, 598–603.

Meirmans PG, Van Tienderen PH (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, **4**, 792–794.

Meraner A, Brandstätter A, Thaler R *et al.* (2008) Molecular phylogeny and population structure of the codling moth (*Cydia pomonella*) in Central Europe: I. Ancient clade splitting revealed by mitochondrial haplotype markers. *Molecular Phylogenetics and Evolution*, **48**, 825–837.

Mochida K (2011) Combination of local selection pressures drives diversity in aposematic signals. *Evolutionary Ecology*, **25**, 1017–1028.

Nei M (1978a) Estimation of average heterozygosity and genetic distance from a number of individuals. *Genetics*, **89**, 538–590.

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

Nokelainen O, Hegna RH, Reudler JH, Lindstedt C, Mappes J (2012) Trade-off between warning signal efficacy and mating success in the wood tiger moth. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 257–265.

Nokelainen O, Lindstedt C, Mappes J (2013) Environment mediated morph linked immune and life history responses in the aposematic wood tiger moth. *Journal of Animal Ecology*, **82**, 653–662.

Nokelainen O, Valkonen J, Lindstedt C, Mappes J (2014) Changes in predator community structure shifts the efficacy of two warning signals in Arctiid moths. *Journal of Animal Ecology*, **83**, 598–605.

O'Hara RB (2005) Comparing the effects of genetic drift and fluctuating selection on genotype frequency changes in the scarlet tiger moth. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 211–217.

Ojala K, Julkunen-Tiitto R, Lindstrom L, Mappes J (2005) Diet affects the immune defence and life-history traits of an Arctiid moth *Parasemia plantaginis*. *Evolutionary Ecology Research*, **7**, 1153.

Ojala K, Lindstrom L, Mappes J (2007) Life-history constraints and warning signal expression in an arctiid moth. *Functional Ecology*, **21**, 1162–1167.

Oostra V, de Jong MA, Invergo BM *et al.* (2011) Translating environmental gradients into discontinuous reaction norms via hormone signalling in a polyphenic butterfly. *Proceedings of the Royal Society B: Biological Sciences*, **278**, 789–797.

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology*, **13**, 55–65.

Peakall R, Smouse PE (2006) GenAlEx 6: genetic analysis in excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Pilot M, DĄBrowski MJ, Jancewicz E, Schtickzelle N, Gliwicz J (2010) Temporally stable genetic variability and dynamic kinship structure in a fluctuating population of the root vole *Microtus oeconomus*. *Molecular Ecology*, **19**, 2800–2812.

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

Raijmann LEL, Menken SBJ (2000) Temporal variation in the genetic structure of host-associated populations of the small ermine moth *Yponomeuta padellus* (Lepidoptera, Yponomeutidae). *Biological Journal of the Linnean Society*, **70**, 555–570.

Ratcliffe JM, Nydam ML (2008) Multimodal warning signals for a multiple predator world. *Nature*, **455**, 96–99.

Raymond M, Rousset F (1995) GENEPOP (Version 3.4) population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

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

Richards-Zawacki CL, Yeager J, Bart HPS (2013) No evidence for differential survival or predation between sympatric color morphs of an aposematic poison frog. *Evolutionary Ecology*, **27**, 783–795.

Robertson KA, Monteiro A (2005) Female *Bicyclus anynana* butterflies choose males on the basis of their dorsal UV-reflective eyespot pupils. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 1541–1546.

Rojas B, Endler JA (2013) Sexual dimorphism and intra-populational colour pattern variation in the aposematic frog *Dendrobates tinctorius*. *Evolutionary Ecology*, **27**, 739–753.

Rosenberg NA (2004) Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

Rousselet J, Zhao R, Argal D *et al.* (2011) The role of topography in structuring the demographic history of the pine processionary moth, *Thaumetopoea pityocampa* (Lepidoptera: Notodontidae). *Journal of Biogeography*, **37**, 1478–1490.

Rowland HM, Ihalainen E, Lindstrom L, Mappes J, Speed MP (2007) Co-mimics have a mutualistic relationship despite unequal defences. *Nature*, **448**, 64–67.

Ruxton GD, Sherratt TN, Speed MP (2004) *Avoiding Attack: The Evolutionary Ecology of Crypsis, Warning Signals and Mimicry*. Oxford University Press, Oxford.

Santos H, Burban C, Rousselet J *et al.* (2011) Incipient allochronic speciation in the pine processionary moth (*Thaumetopoea pityocampa*, Lepidoptera, Notodontidae). *Journal of Evolutionary Biology*, **24**, 146–158.

Saporito RA, Zuercher R, Roberts M, Gerow KG, Donnelly MA (2007) Experimental evidence for aposematism in the dendrobatid poison frog *Oophaga pumilio*. *Copeia*, **4**, 1006–1011.

Siepielski AM, DiBattista JD, Carlson SM (2009) It's about time: the temporal dynamics of phenotypic selection in the wild. *Ecology Letters*, **12**, 1261–1276.

Simmons R (2009) Adaptive coloration and mimicry in: tiger moths and woolly bears: behavior, ecology, and evolution of the Arctiidae Oxford University Press, New York.

Simonato M, Mendel Z, Kerdelhué C *et al.* (2007) Phylogeography of the pine processionary moth *Thaumetopoea wilkinsoni* in the Near East. *Molecular Ecology*, **16**, 2273–2283.

Skipper RA (2009) Revisiting the fisher-wright controversy. *Transactions of the American Philosophical Society*, **99**, 299–322, 346

Stevens M, Ruxton GD (2012) Linking the evolution and form of warning coloration in nature. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 417–426.

Stevens M, PARraga CA, Cuthill IC, Partridge JC, Troscianko TS (2007) Using digital photography to study animal coloration. *Biological Journal of the Linnean Society*, **90**, 211–237.

Symula R, Schulte R, Summers K (2001) Molecular phylogenetic evidence for a mimetic radiation in Peruvian poison frogs supports a Mullerian mimicry hypothesis. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **268**, 2415–2421.

Tamura K, Peterson D, Peterson N *et al.* (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**, 2731–2739.

Thaler R, Brandstatter A, Meraner A *et al.* (2008) Molecular phylogeny and population structure of the codling moth (*Cydia pomonella*) in Central Europe: II. AFLP analysis reflects human-aided local adaptation of a global pest species. *Molecular Phylogenetics and Evolution*, **48**, 838–849.

Todd PA, Ladle RJ, Briers RA, Brunton A (2005) Quantifying two-dimensional dichromatic patterns using a photographic technique: case study on the shore crab (*Carcinus maenas* L.). *Ecological Research*, **20**, 497–501.

Toews DPL, Brelsford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, **21**, 3907–3930.

Valkonen JK, Nokelainen O, Mappes J (2011) Antipredatory function of head shape for vipers and their mimics. *PLoS ONE*, **6**, e22272.

Valkonen JK, Nokelainen O, Niskanen M *et al.* (2012) Variation in predator species abundance can cause variable selection pressure on warning signaling prey. *Ecology and Evolution*, **2**, 1971–1976.

Voudouris CC, Franck P, Olivares J *et al.* (2012) Comparing the genetic structure of codling moth *Cydia pomonella* (L.) from Greece and France: long distance gene-flow in a sedentary pest species. *Bulletin of Entomological Research*, **102**, 185–198.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Weir BS, Hill WG (2002) Estimating F-statistics. *Annual Review of Genetics*, **36**, 721–750.

Wollenberg KC, Lötters S, Mora Ferrer C, Veith M (2008) Disentangling composite colour patterns in a poison frog species. *Biological Journal of the Linnean Society*, **93**, 433–444.

---

J.A.G., O.N. and J.M. conceived the ideas; O.N., R.H. and R.A. collected the data; J.A.G. analysed the data; J.A.G. led the writing; and O.N., J.M. and R.H. contributed.

---

## Data accessibility

All sequence data have been deposited in the European Nucleotide Archive (ENA). Accession numbers

HG764800–HG764812. Data sets used in this study are available at Dryad: doi:10.5061/dryad.s7c52.

## Supporting information

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

**Fig. S1** *Parasemia plantaginis* white morph.

**Fig. S2** Principal components analyses of saturation values in the fore- wing.

**Fig. S3** Forewing of *Parasemia plantaginis* yellow morph.

**Fig. S4** (a) Maximum rate of change in estimating the number of *P. plantaginis* populations (*K*) as inferred by STRUCTURE.

**Table S1** Sampling sites within regions and years of *Parasemia plantagnis* yellow and white morphs.

**Table S2** Analysis of molecular variance (AMOVA) among *Parasemia plantaginis* haplotypes.

**Table S3** Analyses of molecular variance (AMOVA) based on microsatellite allele frequencies of *P. plantaginis* within sampling regions.

**Table S4** Pairwise comparisons between and within *P. plantaginis* morphs across all regions and years calculated from microsatellite data.