INTRODUCTION

Habitat loss and fragmentation represent a fundamental challenge for the conservation of biodiversity (Fahrig, 2003; Haddad et al., 2015). Increased isolation of habitat patches can increase dispersal-related mortality and affect gene flow (Young et al., 1996). Although the genetic effects of habitat fragmentation are not straightforward, increased genetic differentiation is normally expected in species with poor dispersal capacity (e.g., Barluenga et al., 2011; but see Gu et al., 2015). As a result, populations separated by unsuitable
A well-established technique for the detection of the genetic structure is the genotyping-by-sequencing (GBS), which is a reproducible, highly multiplexed next-generation sequencing approach that uses restriction enzymes to reduce genome complexity allowing for simultaneous single nucleotide polymorphism (SNP) discovery and genotyping (Elshire et al., 2011). The major advantages over other available protocols are both technical simplicity (Davey et al., 2011) and that informatics pipelines are publicly available and can be easily adapted to a wide variety of species, either with or without reference genome information (Elshire et al., 2011; Glaubitz...
et al., 2014). GBS, however, has not previously been used for SNP genotyping in any species of the Nymphalidae family. In the current study, we optimized the GBS protocol for the critically endangered marsh fritillary and genotyped a compressive number of individuals from 10 Danish populations. The resulting SNP dataset was used to analyze the genetic structure of the marsh fritillary in order to examine the spatial scale at which genetic differentiation can be detected for a rare, low mobile species.

We aim to answer the following questions (1) Is there evidence of isolation by distance in Danish populations of the marsh fritillary and at what spatial scales can it be detected? (2) Do populations with lower levels of inbreeding also show reduced genetic diversity? and (3) Are the levels of genetic variability and inbreeding related to the genetic isolation among populations?

2 | MATERIALS AND METHODS

2.1 | Sample collection

A total of 300 3rd or 4th instar larvae of the marsh fritillary were collected from 10 sites in Denmark (Figure 1) between August 14 and 31, 2014. At each site, one larva was collected from each of 30 larval clusters. Larvae were kept in separate containers and frozen immediately after collection in the field. The sampling sites were all less than one hectare of mostly homogenous habitat. The geographical coordinates of individual larval clusters within sites were not considered for this study. The locations were numbered as (1) Bruså, (2) Tranum Skydetræn, (3) Overklitten Sø, (4) Sandmosen, (5) Vågholt Mose, (6) Troldekærvej, (7) Knasborgvej, (8) Videsletengen, (9) Milrimvej, and (10) Strandby.

2.2 | DNA extraction and genotyping-by-sequencing (GBS) protocol optimization

The total genomic DNA was extracted from larvae using the DNeasy® Blood & Tissue Kit (Qiagen Inc., Hilden, Germany) according to the manufacturer’s protocol for purification of total DNA from insects. DNA quantity and quality was verified using a fluorometer (Qubit®, Thermo Fisher Scientific Inc.) and by running 100 ng of each DNA sample on a 1% agarose gel, respectively.

For optimization of the standard GBS protocol (Elshire et al., 2011) for E. aurinia, a single DNA sample (400 ng) was digested for 2 hr with the restriction enzymes ApeKI, EcoT22I, and PstI, in separate essays, using a tenfold excess of enzyme and reaction conditions as specified by the endonuclease manufacturer (New England Biolabs). After ligation of appropriate adapters (adapter amounts were determined by titration as described in REF(Elshire et al., 2011) and PCR (see below)), fragment size distributions of each
| Region     | Population Name | ID          | Latitude      | Longitude     | Parameter | N   | P%   | Ne   | H₀   | Hₑ   | uHₑ  | Fₑ      |
|------------|-----------------|-------------|---------------|---------------|-----------|-----|------|------|------|------|-------|---------|
| Region A   | Bruså            | Pop1        | 56°55.519'N  | 9°26.312'E    | Mean      | 30  | 64.47| 1.343| 0.183| 0.203| 0.207 | 0.100   |
|            |                 |             |               |               | SE        |     |      | 0.021| 0.011| 0.011| 0.011 | 0.019   |
| Region B   | Tranum Skydeteraen | Pop2       | 57°10.948'N  | 9°30.338'E    | Mean      | 30  | 81.13| 1.368| 0.194| 0.228| 0.232 | 0.141   |
|            |                 |             |               |               | SE        |     |      | 0.018| 0.009| 0.010| 0.010 | 0.018   |
|            | Overklitten Sø  | Pop3        | 57°11.000'N  | 9°36.604'E    | Mean      | 28  | 76.10| 1.372| 0.190| 0.224| 0.228 | 0.130   |
|            |                 |             |               |               | SE        |     |      | 0.020| 0.010| 0.010| 0.011 | 0.018   |
|            | Sandmosen       | Pop4        | 57°33.647'N  | 10°15.496'E   | Mean      | 28  | 72.33| 1.365| 0.188| 0.221| 0.225 | 0.126   |
|            |                 |             |               |               | SE        |     |      | 0.020| 0.010| 0.010| 0.011 | 0.017   |
| Region C   | Vågholt Mose    | Pop5        | 57°34.596'N  | 10°24.516'E   | Mean      | 25  | 85.22| 1.376| 0.189| 0.237| 0.242 | 0.179   |
|            |                 |             |               |               | SE        |     |      | 0.017| 0.009| 0.009| 0.009 | 0.019   |
|            | Troldkærvej     | Pop6        | 57°33.899'N  | 10°24.938'E   | Mean      | 17  | 74.21| 1.351| 0.179| 0.216| 0.224 | 0.137   |
|            |                 |             |               |               | SE        |     |      | 0.019| 0.010| 0.010| 0.010 | 0.021   |
|            | Knasborgvej     | Pop7        | 57°29.112'N  | 10°30.345'E   | Mean      | 30  | 85.85| 1.367| 0.190| 0.231| 0.235 | 0.142   |
|            |                 |             |               |               | SE        |     |      | 0.018| 0.009| 0.009| 0.009 | 0.017   |
|            | Videsletengen   | Pop8        | 56°55.519'N  | 9°26.312'E    | Mean      | 30  | 90.25| 1.380| 0.188| 0.241| 0.246 | 0.185   |
|            |                 |             |               |               | SE        |     |      | 0.017| 0.008| 0.009| 0.009 | 0.018   |
|            | Milrimvej       | Pop9        | 57°10.275'N  | 9°32.226'E    | Mean      | 27  | 88.99| 1.364| 0.171| 0.232| 0.237 | 0.225   |
|            |                 |             |               |               | SE        |     |      | 0.017| 0.008| 0.009| 0.009 | 0.019   |
|            | Strandby        | Pop10       | 57°11.000'N  | 9°36.604'E    | Mean      | 28  | 84.28| 1.362| 0.170| 0.225| 0.229 | 0.208   |
|            |                 |             |               |               | SE        |     |      | 0.018| 0.008| 0.010| 0.010 | 0.018   |

Note: The region, population name, id, latitude, and longitude are given for reference. The number of individuals sampled in each population (n) and proportion of polymorphic loci (P%) are presented along with mean and standard error (SE) of mean effective alleles (Ne) based on 318 SNPs, observed heterozygosity (H₀), expected heterozygosity (Hₑ), unbiased expected heterozygosity (uHₑ), and inbreeding coefficient (Fₑ).
test library were visualized using an Experion (Bio-Rad, Hercules, California, USA) (Figure S1). Based on these results, we selected the libraries obtained from the EcoT22I digests to maximize sequence coverage from GBS.

### 2.3 Preparation of Illumina libraries for next-generation sequencing

Three 96-plex EcoT22I GBS libraries, comprising 285 DNA samples and three negative (no DNA) control, were prepared according to Elshire et al. (2011). Low DNA concentration samples (n = 15) were discarded and not submitted for sequencing. Briefly, individual DNA samples were digested with the restriction enzyme and adapters were ligated as described previously. The adapters comprised a set of 96 different barcodes containing adapters and a “common” adapter. Individual ligations were pooled, and purified using QiaQuick PCR purification kit (Qiagen). Genomic fragments were then amplified in a 50-μl volume containing 2-μl DNA fragments, 1× Taq Master Mix (New England Biolabs), and 25 pmol, each, of the following primers: (a) 5′-AATGATACGGCGACCACCGAGATCTACACT CTTTCCCTAACACGACGCTTCTTCCGATCT-3′ and (b) 5′-CAAGCAG AAGACGCGTATGACAGAGTTCGCTTCTCGGTCGACGCACG TCTTCCGATCT-3′. PCR cycling consisted of 72°C for 5 min, 98°C for 30 s, followed by 18 cycles of 98°C for 30 s, 65°C for 30 s, and 72°C for 30 s, with a final extension step at 72°C for 5 min. The EcoT22I GBS library was purified again, as above, and an aliquot was run on an Agilent BioAnalyzer 2100 for evaluation of fragment sizes and the presence of adapter dimers. After quantification on a Nanodrop 2000 (Thermo Scientific), the three 96-plex libraries obtained from the Illumina HiSeq 2000. GBS library preparation (Elshire et al., 2011), sequencing, and SNP calling were performed at the Genomic Diversity Facility (GDF) at Cornell University’s Biotechnology Resource Center.

### 2.4 DNA sequence analysis: SNP discovery and genotyping

Illumina raw reads were processed using the default parameter of the Universal Network-Enabled Analysis Kit (UNEAK) (Lu et al., 2013) for species without a reference genome. This pipeline was implemented in TASSEL version 3.0.166 (Glaub et al., 2014) and used for tag alignment and subsequent SNP calling. Briefly, the raw Illumina DNA sequence data (100-bp qseq files) were first trimmed to remove barcodes. The sequence remnants were then either trimmed further or padded with 3’ A’s to 64-bp lengths. Sequences were then aligned to each other, both to identify unique sequences, or “sequence tags”, and to generate clusters of related sequences. For each cluster, a network was generated, in which sequence tags were organized according to mutation steps (i.e., mutational relationship). A single base-pair mismatch was allowed among cluster members. Networks were then filtered such that only SNPs originating from reciprocal tag pairs were retained (see Lu et al., 2013). SNPs from more complicated networks that often result from alignment of paralogs and repeats, or sequencing errors were discarded. To further reduce the impact of sequencing errors, we also set the error tolerance rate (ETR) parameter to 0.03, slightly below the expected Illumina sequencing error rate (0.04%). Failed samples (nonblank), defined as those with less than 10% of the mean reads per sample coming from the lane on which they were sequenced, were discarded (n = 5).

The resulting raw SNP dataset from the UNEAK pipeline was further filtered using Golden Helix SNP and Variation Suite (SVS version 7.2.2, Golden Helix, Bozeman, MT) and PLINK v1.07 (Purcell et al., 2007) softwares. First, the dataset was filtered by the application of genotype-level filters to remove genotypes with low read depths (RD) and/or low genotype quality (GQ). Thus, genotypes with RD ≤ 4x and GQ ≤ 98 were considered as missing. Later, we removed all SNPs and individuals with call rates <80%. In addition, SNPs with a minor allele frequency (MAF) < 0.05 were removed. Loci with a mean-observed heterozygosity > 0.6 were also discarded to filter out potential paralogs. The SNP set was then pruned for linkage disequilibrium (LD) by excluding markers in strong LD (pairwise genotype correlation r² > .5) in a window of 50 SNPs (sliding window overlap 10 SNPs at a time). This filtering process is described in Figure S2.

### 2.5 Genetic variability and population structure

Genetic variability in each population was assessed by the calculation of observed heterozygosity (H_o), expected heterozygosity (H_e), unbiased heterozygosity (uH_e) and inbreeding coefficient (Fis), mean percentage polymorphic loci (%P), and mean effective alleles (Ne) were likewise calculated using GenAIEx 6.501 (Peakall & Smouse, 2012). Deviations from Hardy-Weinberg Equilibrium (HWE) (probability test) were analyzed using GENEPOP v4.3 on all the populations pooled together and on each population considered singularly (Rousset, 2008).

Population pairwise Fst were calculated to reveal the genetic differentiation using GenAIEx 6.501, and Fisher’s exact probability test for testing for genetic differentiation was carried out using GENEPOP v4.3. For every population, the mean of pairwise Fst values between the population and all other populations was considered as an index of isolation.

A maximum likelihood-based clustering algorithm implemented in ADMIXTURE v1.23 (Alexander et al., 2009) was applied to the entire dataset to identify the putative ancestral cluster(s) within the samples as well as to assess the extent of genetic admixture. Clustering was performed 100 times for all K-values from K = 2 to K = 12, and the best-fitting K was selected based on the lowest cross-validation error (CVE). Additionally, a principal component analysis (PCA) was performed in R (R Development Core Team https://www.r-project.org/), based on the genetic distances using GenAIEx v6.501 to examine how populations would cluster along principal component axes 1 and 2. Finally, Mantel’s tests were performed.
on all the populations investigated in order to determine whether there was an overall correlation between geographic distance and genetic divergence (Smouse et al., 1986). Some authors argue that spatial structures in the dataset can enhance isolation by distance (Legendre et al., 2015). To control for such a potential effect, we ran one test based on all pairwise comparisons and one test on a subset of pairwise comparisons excluding pairs from two different regions.

3 | RESULTS

The UNEAK pipeline recovered 30,137 bi-allelic SNP loci (n = 280; 5 samples failed the UNEAK pipeline). However, most of these SNPs had low coverage or were only present in a small number of individuals. After the complete filtering procedure, 318 SNPs were maintained in our matrix for 273 individuals with an overall call rate of 93.57% (see Figure S2). Over all of these loci, the mean coverage per locus per individual was 87.88 (max coverage per individual 207.21, min 8.21).

Deviations from HWE were found to be highly significant for all the 10 populations investigated (p < .001 in all cases), both when pooled altogether and when considered singularly. All the deviations were due to heterozygote deficiency as can be seen by the positive $F_{IS}$ values ranging from 0.1 to 0.228 (Table 1). Genetic variability parameters, observed heterozygosity ($H_o$), unbiased heterozygosity ($uH_o$) and inbreeding coefficient ($F_{IS}$), mean percentage polymorphic loci (%P), and mean effective alleles (Ne), are listed in Table 1. Genetic divergence between populations ranged from 0.028 to 0.1 (Table 2). All the pairwise $F_{ST}$ values were highly significant (p < .001).

Populations were strongly clustered into three geographic regions; the minimum CVE value in the ADMIXTURE analysis suggested an optimal number of genotypic clusters for $K = 6$ (CVE = 0.525) (Figure S3). The graphical visualization of the ADMIXTURE results for the 273 individuals and K ranging from two to 12 clusters is shown in Figure 2. When $K = 6$, E. aurinia is clearly subdivided into different genetic clusters that mainly corresponded with the three geographic regions of the species distribution (i.e., A, B, C) and each of the sampled populations: Region A (population 1) is characterized by a single genetic cluster; Region B is subdivided into two different clusters, one including populations 2 and 3 (dark blue) and an additional cluster exclusively including population 4 (light blue); and Region C genetic subdivision is more complex and characterized by 3 private clusters (red, green and orange clusters). However, while population 10 is mainly characterized by a single genetic component (orange cluster), populations 5 to 9 showed varying levels of genetic admixture of red and green clusters, with populations 5, 7, and 8 showing a predominance of red cluster and 6 and 9 populations a predominance of the green genetic component. Overall, ADMIXTURE results for $K ≥ 6$ provided similar outputs, but progressively increased the level of genetic resolution within the E. aurinia populations, with $K = 10$ providing a genetic clustering result mainly differentiating each of the sampled marsh fritillary populations with varying levels of genetic admixture (Figure 2).

Also, the PCA analysis detected three distinct clusters which were clearly separated by PC1 and PC2 (Figure 3), and they coincide exactly with the three regions A, B, and C of the distribution of the species in Denmark (Figure 1). The genetic relationship among the 10 populations quantified using a principal component analysis (PCA) in which the first two axes (PC1 and PC2) explained 5.09% and 3.27% of the variation, respectively. The first cluster includes only population 1 (region A), the second cluster includes populations 2, 3, and 4 (region B), and the third cluster includes populations 5-10 (region C).

Strong evidence for isolation by distance was found across populations; the Mantel tests were found to be highly significant (least square regression analyses; p < .001). The regression of genetic divergence ($F_{ST}$) against log_{10} of geographic distance in km was highly significant for all pairwise population comparisons ($R^2 = .73$, df = 43, p < .001; Figure 4a) as well as for the subset excluding comparisons among populations in each of the three regions in which the marsh fritillary occurs ($R^2 = .79$, df = 22, p < .001; Figure 4b).

### Table 2: Pairwise $F_{ST}$ values (lower left) and geographical distance in kilometers (upper right) among all pairs of the 10 Danish populations of the marsh fritillary (Euphydryas aurinia) were investigated

| Pop1   | Pop2     | Pop3     | Pop4     | Pop5     | Pop6     | Pop7     | Pop8     | Pop9     | Pop10    |
|--------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| Pop1   | 28.9     | 28.0     | 30.5     | 86.3     | 94.1     | 93.2     | 92.5     | 90.3     | 89.7     |
| Pop2   | 0.083    | 2.3      | 6.3      | 61.8     | 69.6     | 69.8     | 69.3     | 67.3     | 69.0     |
| Pop3   | 0.092    | 0.036    | 4.6      | 61.3     | 69.1     | 69.2     | 68.7     | 66.6     | 68.0     |
| Pop4   | 0.091    | 0.049    | 0.048    | 57.3     | 65.1     | 65.0     | 64.4     | 62.3     | 63.5     |
| Pop5   | 0.086    | 0.065    | 0.074    | 0.078    | 7.8      | 9.2      | 9.4      | 8.2      | 17.0     |
| Pop6   | 0.100    | 0.084    | 0.098    | 0.092    | 0.052    | 5.5      | 6.7      | 7.6      | 17.1     |
| Pop7   | 0.088    | 0.067    | 0.074    | 0.083    | 0.034    | 0.053    | 1.4      | 3.2      | 11.7     |
| Pop8   | 0.079    | 0.063    | 0.071    | 0.074    | 0.03     | 0.044    | 0.028    | 2.2      | 10.4     |
| Pop9   | 0.082    | 0.066    | 0.072    | 0.077    | 0.037    | 0.051    | 0.041    | 0.031    | 9.8      |
| Pop10  | 0.096    | 0.069    | 0.078    | 0.079    | 0.048    | 0.062    | 0.048    | 0.040    | 0.050    |

Note: All $F_{ST}$ comparisons were highly significant (p < .001).
4 | DISCUSSION

We used a genotyping-by-sequencing (GBS) approach to increase the knowledge on the population genetic structure of the critically endangered marsh fritillary butterfly in Denmark. This study documents the identification of an informative SNP loci panel and demonstrates that the GBS approach represents a powerful tool to define genetic relationships at the intraspecific level. Moreover, this SNP panel provides an important genetic resource for further genetic studies of the marsh fritillary and is a cost-effective and rapid method that can well describe the genetic variability of other non-model species with limited genetic resources.

The stringent filtering procedure that we have applied (GQ = 98) (compared with most studies where the GQ filter, is more commonly

**FIGURE 2** ADMIXTURE analysis for ancestry population clusters (K = 2–12) among the 10 analyzed populations of the marsh fritillary (*Euphydryas aurinia*) (*n* = 273) based on 318 SNPs. Populations are separated by white vertical lines

**FIGURE 3** Plot of principal component axes 1 and 2 based on a principal component analysis of the relationship among populations based on genetic distance among 10 populations of the marsh fritillary (*Euphydryas aurinia*). The first PCA axis explained 5.09%, and the second PCA axis explained an additional 3.27% of the variation in the data. Colors and symbols are combined to maximize readability (circles are for Region 1, squares for Region 2, and triangles for Region 3). The sampled populations are listed as in Figure 1: (1) Bruså, (2) Tranum Skydertann, (3) Overklitten Sø, (4) Sandmosen, (5) Vågholt Mose, (6) Troldkærvej, (7) Knasborgvej, (8) Videsletengen, (9) Milrimvej, and (10) Strandby
Among the three main regions where the marsh fritillary is found, even when accounting for the spatial structure in our data by omitting comparisons among the three regions in which the marsh fritillary is found (Legendre et al., 2015). Detailed capture-mark-recapture studies have demonstrated that most dispersal events are shorter than one kilometer and, only in rare cases, marsh fritillary butterflies disperse more than five kilometers (Johansson et al., 2019; Zimmermann et al., 2011). This suggests that under current levels of fragmentation, isolation by distance can be detected at the same spatial scale as the dispersal capacity of the species. Most of the clusters identified by the ADMIXTURE showed well-defined genetic clusters coinciding with distinct geographic regions of the distribution of the species in Denmark.

The results of ADMIXTURE show that the populations in regions A and B are more genetically similar while the PCA plot shows a tighter genetic similarity between populations in regions B and C. However, the common approach for detecting the number and subdivision of clusters with the use of ADMIXTURE assumes equilibrium in the genetic conditions, no deviation from HWE, and no linkage disequilibrium within a cluster. Such assumptions are clearly violated in our sample, whereas PCA analyses are not biased by such deviations from genetic equilibrium. Therefore, a cautionary approach should be undertaken when interpreting the genetic relatedness between populations and also the $F_{ST}$ distances can mislead as that this genetic-distance estimator is also based on the assumption of genetic equilibrium within the populations compared. The strong negative relationship found between mean $F_{ST}$ and $F_{IS}$ provided further evidence for the negative consequences of fragmentation on the genetic variability and the in-breeding level within populations of the marsh fritillary in Denmark.

Our results suggest that further efforts are needed to maintain genetic diversity in this species in Denmark. Source-sink dynamics will effectively increase mortality from isolated populations because dispersing individuals will be unable to locate suitable habitat. In addition, reduced mixing of populations will affect genetic diversity and ultimately cause inbreeding (Pertoldi et al., 2007; Sigaard et al., 2008). A breeding program for the marsh fritillary in the United Kingdom demonstrated a strong positive effect on reproduction by mixing populations from Cumbria and Scotland into a hybrid stock (Porter & Ellis, 2011). Since all populations in our study from Denmark were inbred, similar positive effects of mixing populations in Denmark could be expected. Although several strongholds for the marsh fritillary exist in Denmark, establishing a breeding program could be relevant. Hybrid populations could serve as a way to secure genetic diversity and may be relevant for potentially translocating the species to unoccupied regions of Denmark with suitable habitat networks (Brunbjerg et al., 2017).

Population structure has been studied in many butterfly species using microsatellite markers (Saccheri et al., 2004; Smee et al., 2013; Vandewoestijne et al., 2011; Zeisset et al., 2005). However, due to the formidable challenges involved in developing informative microsatellite markers (Nève & Meglécz, 2000), most studies have relied on a small number of markers with limited resolution (Sigaard et al., 2008). Methods like GBS can help guide priorities for other species in similar situations (Elshire et al., 2011).

We found isolation by distance for patches across a wide range of distances among populations. We expected that populations in the three regions investigated to be completely isolated from each other, but the correlation between geographic distance and genetic divergence suggests that some gene flow may occur among these regions. The genetic differentiation observed among the three regions is, however, compatible with the expectation in species with poor dispersal capacity. Given the fact that the populations are
separated by unsuitable habitats, we expect that only very rare dispersal events from one region to another is occurring. Long and relatively rare dispersal events have been detected among populations of marsh fritillary in the Czech Republic populations (Zimmermann et al., 2011). Further studies are needed to elucidate how variation in habitat characteristics like resource availability in different life stages affect propensity for dispersal and to demonstrate that such long-distance dispersal among regions does occur. In order to mitigate further losses of genetic diversity, conservation efforts targeting rare species like the marsh fritillary existing in fragmented landscapes like in Denmark should concentrate on further enhancing connectivity among existing habitat patches.

ACKNOWLEDGMENTS
This work was supported by the Aalborg Zoo Conservation Foundation (AZCF) and 15. Juni Fonden (grant number: 2014-A-13 and 2017-N-6). The Danish Nature Agency is acknowledged for permission to collect marsh fritillary larvae.

CONFLICT OF INTEREST
The authors have no conflict of interest.

AUTHOR CONTRIBUTIONS
Cino Pertoldi: Conceptualization (lead); Data curation (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Methodology (lead); Project administration (equal); Resources (equal); Software (equal); Supervision (equal); Validation (equal); Writing-original draft (lead); Writing-review & editing (equal).

Aritz Ruiz-Gonzalez: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (equal); Resources (equal); Software (equal); Supervision (equal); Validation (equal); Writing-original draft (equal); Writing-review & editing (equal).

Simón Bahndorff: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (equal); Resources (equal); Software (equal); Supervision (equal); Validation (equal); Visualization (equal); Writing-original draft (equal); Writing-review & editing (equal).

Nanna Renee Lauridsen: Conceptualization (supporting); Data curation (supporting); Formal analysis (supporting).

Thøger Nisbeth Henriksen: Conceptualization (supporting); Data curation (supporting). Annette Eskildsen: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (equal); Resources (equal); Software (equal); Supervision (equal); Validation (equal); Writing-review & editing (equal).

Toke Thomas Høye: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (equal); Resources (equal); Software (equal); Supervision (equal); Validation (equal); Visualization (equal); Writing-original draft (equal); Writing-review & editing (equal).

DATA AVAILABILITY STATEMENT
Data for this study are available at Dryad: https://datadryad.org/stash/share/XetcrVN2qTVd3aRg1eK4LxAmZXXBegZtm0s kGawIr1o.

ORCID
Cino Pertoldi https://orcid.org/0000-0002-4644-8981
Toke Thomas Høye https://orcid.org/0000-0001-5387-3284

REFERENCES
Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Research, 19(9), 1655–1664. https://doi.org/10.1101/gr.094052.109

Asbirk, S., & Christensen, T. (2000). Handlingsplan for bevarening af den truede sommerfugl hedeplevning Euphydryas aurinia. http://www.sns.dk/natur/bevaring/pdf/pdf/ hedeplet.pdf

Barluenga, M., Austerlitz, F., Elzinga, J. A., Teixeira, S., Goulet, J., & Bernasconi, G. (2011). Fine-scale spatial genetic structure and gene dispersal in Silene latifolia. Heredity, 106(1), 13–24. https://doi.org/10.1038/hdy.2010.38

Brumberg, A. K., Høye, T. T., Eskildsen, A., Nygaard, B., Damgaard, C., & Ermans, R. (2017). The collapse of marsh fritillary (Euphydryas aurinia) populations associated with declining host plant abundance. Biological Conservation, 211, 117–124. https://doi.org/10.1016/j.biocon.2017.05.015

Caballero, A. (1994). Developments in the prediction of effective population size. Heredity, 73, 657–679. https://doi.org/10.1038/hdy.1994.174

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nature Reviews Genetics, 12(7), 499–510. https://doi.org/10.1038/nrg3012

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One, 6(5), e19379. https://doi.org/10.1371/journal.pone.0019379

Eskildsen, A., Carvalheiro, L. G., Kissling, W. D., Biesmeijer, J. C., Schweiger, O., & Høye, T. T. (2015). Ecological specialization matters: Long-term trends in butterfly species richness and assemblage composition depend on multiple functional traits. Diversity and Distributions, 21(7), 792–802. https://doi.org/10.1111/dad.12340

Fahrig, L. (2003). Effects of habitat fragmentation on biodiversity. Annual Review of Ecology Evolution and Systematics, 34, 487–515. https://doi.org/10.1146/annurev.ecolsys.34.011802.132419

Fraser, D. J., Debes, P. V., Bernatchez, L., & Hutchings, J. A. (2014). Population size, habitat fragmentation, and the nature of adaptive variation in a stream fish. Proceedings of the Royal Society B: Biological Sciences, 281(1790), 20140370. https://doi.org/10.1098/rspb.2014.0370

Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., & Buckler, E. S. (2014). TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. PLoS One, 9(2), e90346. https://doi.org/10.1371/journal.pone.0090346

Gu, Q. H., Zhou, C. J., Cheng, Q. Q., Li, X. J., Zhu, G. R., Zhang, M., & Dong, J. (2015). The perplexing population genetic structure of Bellamyra purificata (Gastropoda: Viviparidae): Low genetic differentiation despite low dispersal ability. Journal of Molluscan Studies, 81, 466–475. https://doi.org/10.1093/mollus/eyv017
Zimmermann, K., Fric, Z., Jiskra, P., Kopeckova, M., Vlasanek, P., Zapletal, M., & Konvicka, M. (2011). Mark-recapture on large spatial scale reveals long distance dispersal in the Marsh Fritillary, Euphydryas Aurinia. *Ecological Entomology*, 36(4), 499–510. https://doi.org/10.1111/j.1365-2311.2011.01293.x

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Pertoldi, C., Ruiz-Gonzalez, A., Bahrndorff, S., Renee Lauridsen, N., Nisbeth Henriksen, T., Eskildsen, A., & Høye, T. T. (2021). Strong isolation by distance among local populations of an endangered butterfly species (*Euphydryas aurinia*). *Ecology and Evolution*, 11, 12790–12800. https://doi.org/10.1002/ece3.8027