Short title: Black-veined white extirpation

The last days of *Aporia crataegi* (L.) in Britain: evaluating genomic erosion in an extirpated butterfly

Rebecca Whitla[1], Korneel Hens[1,2], Casper Breuker[1,2], Timothy G. Shreeve[1], Saad Arif[1,2,*]

[1] Department of Biological and Medical Sciences, Oxford Brookes University, OX3 0BP, United Kingdom

[2] Centre for Functional Genomics, Oxford Brookes University, OX3 0BP Oxford, United Kingdom

* Author for correspondence

Keywords: *Aporia crataegi*, inbreeding, genetic load, historical DNA, genomic erosion, genetic diversity

Correspondence:

Saad Arif

Centre for Functional Genomics, Oxford Brookes University, OX3 0BP Oxford, United Kingdom

Department of Biological and Medical Sciences, Oxford Brookes University, OX3 0BP, United Kingdom

Tel: +44 (0)1865 483437

Email: sarif@brookes.ac.uk

Abstract

Current rates of climate change and habitat degradation are causing biodiversity declines globally, with global extinction rates far beyond estimated background extinction rates. Studies on vertebrates highlight how conservation genomics can be an effective tool to identify and manage threatened populations. However, it is unclear how metrics of genomic erosion, commonly applied in conservation genomics, translate to invertebrate species which can have markedly different population sizes and life history strategies. The Black-veined White butterfly (*Aporia crataegi)* was extirpated from Britain in the 1920s. Here, we sequenced historical DNA from 17 museum specimens collected between 1854 and 1924 to determine levels of genomic erosion in the species as it approached extirpation. We compare these results to those from the Common Blue butterfly (*Polyommatus icarus*) - a species with stable demographic trends in Great Britain. We provide evidence for a bottleneck of British butterflies around the period of post-glacial colonisation of the British Isles. Despite a 4X difference in effective population sizes around the time of post-glacial colonisation, reduction in genome-wide heterozygosity (our metric for genetic diversity) in *A. crataegi* (17.9%) was comparable to that in *P. icarus* (14.5%). More symptomatic of *A. crataegi*'s disappearance were marked increases in size and proportion of runs-of-homozygosity (RoH) and accumulation of putatively mildly and weakly deleterious variants. Our results support the notion that metrics with more direct links to host fitness and function, such as inbreeding and genetic load, may be more informative of population decline and thus more valuable in identifying vulnerable insect populations.

Introduction

Global extinction rates have increased over the last 500 years (Barnosky *et al*., 2011, Ceballos *et al*., 2015, Cowie *et al*., 2022). Rates have accelerated over the last 100 years and *circa* 25% of species recently assessed by the IUCN are now threatened with extinction (IPBES, 2019). In the UK, *c.* 41% of assessed flora and fauna species (8,431) have declined in abundance since the 1970s, and *c.* 2% of all species have gone extinct since the 1500s (Hayhow *et al*., 2019). Invertebrates count for 97% of all animals globally, but assessments for many invertebrates are data deficient, precluding accurate assessment of their status (Karam-Gemael *et al*., 2020). Reasons for insufficient data include underfunding for less charismatic and economically important groups (Cardoso and Leather, 2019) and difficulties in sampling and identification (Wagner *et al*., 2021).

Invertebrates, including insects, are a major component of terrestrial biodiversity with important roles in ecosystem functioning. Among insects, butterflies are one of the best-studied groups (for example, Warren, 1992; Thomas, 2005; Dincă *et al*., 2015; Middleton-Welling *et al*., 2020; Warren *et al*., 2021), used as bioindicators of environmental and climate change (Dennis *et al*., 2003; Hill *et al*., 2021) and serve as an important invertebrate model for communicating science to society (Preston *et al*., 2021). Since the 1970s, standardised and systematic monitoring of butterflies across Europe has provided a comprehensive record of changes in biodiversity. Compared to most of continental Europe, a large proportion of UK butterflies are now either regionally extinct (4 of 62 species) or considered as critically endangered or vulnerable (19/62) (Fox *et al*., 2022).

By quantifying genomic erosion, conservation genomics can be a valuable tool in conservation, helping identify vulnerable populations and informing translocations and/or breeding programmes (Díez-del-Molino *et al.,* 2018; Bosse and van Loon, 2022). Metrics for genomic erosion, associated with reductions of fitness (Bosse and van Loon, 2022), aim to estimate genetic diversity, effective population size (N_e), levels of inbreeding and genetic load. All have a strong theoretical basis to be correlated with and potentially exacerbating extinction risk in small and vulnerable populations (Gomulkiewicz and Holt, 1995; Charlesworth and Willis, 2009; Lynch *et al*., 1995a). Negative impact of inbreeding depression in small populations is well established (Saccheri *et al*., 1998; Mattila *et al*.,

2012; Neaves *et al*., 2015; Rowe *et al.,* 2017) but the utility of metrics of genomic erosion, particularly those relying solely on putatively neutral genetic markers or diversity, in predicting extinction risk is currently under debate (Kardos *et al*., 2016; Teixeira and Huber, 2021; Dussex *et al*., 2023). A more a nuanced understanding is emerging from the accumulation of population genomics of small and threatened populations using whole genome resequencing data to estimate multiple metrics of genomic erosion simultaneously (DeWoody *et al*., 2021; Bosse and van Loon, 2022; Dussex *et al*., 2023), but is largely limited to examples from terrestrial vertebrates and our understanding of genomic erosion in insects is far more limited (Webster *et al*., 2023; but see de-Dios *et al*., 2023). Comparative analyses of genomic erosion on endangered, but also stable and expanding insect populations, could prove extremely useful for detecting insect declines or identifying vulnerable species or populations from small or limited samples.

Here, we examine trends of genomic erosion in *Aporia crataegi* (L.) (Black-veined White butterfly), which was extirpated in the British Isles in the mid-1920s (Pratt, 1983), contrasting these with extant populations in continental Europe. We also provide a comparison between British and continental European populations of *Polyommatus icarus* (Rott.) (Common Blue butterfly) – a species that has exhibited no change in long term population trends (https://www.gov.uk/government/statistics/butterflies-in-the-wider-countryside-uk). *A. crataegi* is widespread throughout much of Europe, Asia and north Africa. It has recently been lost from the Netherlands, Czechia (van Swaay *et al*., 2010), and South Korea (Kim *et al*., 2015), and is declining in other European countries (van Swaay *et al*., 2010). Within Britain, it was restricted to southern and central England and Wales. The species was extirpated in Britain *c.* 1925 but there is evidence that populations were in decline from the mid to late 19th century (Dale, 1887; South 1906; Pratt, 1983) and was primarily only found in southeast England by the end of the 19th Century. Its decline in Europe is typically attributed to habitat loss, however this does not seem to be the case for its extirpation in Britain (McLachlan, 1893). The exact reason for its extinction in Britain is unknown, but climate variability (particularly high September rainfall and extreme winters) along with pesticide use, parasites, microbial or fungal pathogens and predation by small birds all have been suggested as potential reasons (reviewed in Pratt, 1983). Numerous “accidental” releases, recorded as early as the

19th century (Pratt, 1983), from butterfly enthusiasts to reintroduce the species from continental European stock have been made but none have been successful.

To examine whether *A. crataegi* in Britain exhibited any symptoms of genomic erosion during its extirpation process, we sequenced historical specimens of *A. crataegi* from Britain and continental Europe dating from 1854-1924. We specifically addressed whether the British *A. crataegi* were distinct from continental Europe based on genetic structuring and demographic history over the last 10 kilo years ago (Kya). We then examined individual heterozygosity, runs-of-homozygosity (RoH, a genomic measure of recent inbreeding), and genetic load to explain the vulnerability of the British Isles populations prior to extinction and contextualise these finding by comparison with corresponding analyses of British and continental European *P. icarus*, a species with stable long-term population trends.

Materials and Methods

Historical* A. crataegi *Specimen Selection

Seventeen *A. crataegi* specimens from Britain (GB), Belgium, and France, (Fig. 1A, Table S1) were sampled from Oxford University Museum of Natural History (OUNHM), Oxford, England and the Natural History Museum (NHM), London, England. Sampling dates ranged from 1854-1924. For comparison, we only used museum samples of Continental populations to avoid any potential bias stemming from the damage to the historical DNA. Up to two legs per specimen were taken from pinned samples for DNA extraction, although in some cases only a single leg was available.

Extraction, Library Preparation and Sequencing of Historical DNA

DNA extractions were undertaken on a laboratory bench used only for historical DNA (hDNA) extractions to prevent contamination from modern DNA. Extraction blanks were prepared alongside samples to monitor potential cross-contamination. Legs of specimens were fully macerated using a Tissuelyser II bead mill (Qiagen) and extracted using a QIAamp DNA Micro kit (Qiagen) using the

standard protocol with a final elution volume of 20µL. Total DNA yield was measured using a Qubit 4 Fluorometer (Life Technologies) with a Qubit dsDNA HS Assay Kit. Eluted DNA was stored in Eppendorf DNA LoBind Tubes at -20°C until library preparation.

Genomic library preparation of hDNA was performed using the NEBNext Ultra II DNA Library Prep Kit for Illumina. A CUT&RUN (Liu *et al.,* 2018) protocol was used, modified for use with low yield samples and samples with short fragment sizes. NEBNext® Dual Index Multiplex Oligos for Illumina (Set 1) Primers were used constructing multiplexed libraries. SPRIselect beads (Beckman Coulter) were used for size exclusion post primer ligation at 1.75X or 2.2X concentration.

The number of PCR cycles used for amplification was dependent on the input DNA quantity; 12 cycles for samples < 5ng and 10 cycles for samples ≥ 5ng. Libraries were then purified with 1.2X SPRIselect beads to remove adapter dimer. Library enrichment was performed using either a Q5U® Hot Start High-Fidelity DNA Polymerase (New England Biolabs) or the NEBNext Ultra II Q5DNA polymerase (Table S1). We did not attempt to repair the hDNA template prior to library preparation. However, the NEBNext Ultra II DNA Library Prep Kit makes use of USER enzyme during the adapter ligation step. This excises Uracil bases from the ligated adapters and is required when using NEBNext adapters, but in case of hDNA template with postmortem deamination damage, the enzyme will also excise Uracil bases on the 5’ end of the hDNA template. With subsequent paired-end Illumina sequencing this results in partial repair with 5’ ends exhibiting little to no signs of historical damage whereas the damage profile on 3’ ends reflects typical postmortem DNA damage.

Quality control for the libraries was performed using a Qubit 4 Fluorometer (Life Technologies) and an Agilent Bioanalyser, using the Agilent High Sensitivity DNA analysis Chip Kit. Sample libraries were pooled equimolarly, with extraction blanks pooled at a 1:10 molar ratio in proportion to sample libraries. Sequencing was performed using the NovaSeq Illumina platform using a 150 base pair (bp) paired-end (150PE) strategy at Novogene (Cambridge).

***Sampling Modern specimens of* P. icarus**

Four individuals of *P. icarus* were collected from across the British Isles and an additional three from Averyon, France (Table S2, Fig. 1A) and DNA extraction is described Arif *et al.* (2021). DNA was normalised to 10 ng/ul and library preparation and sequencing, to a theoretical coverage of 15-50X (Table S2), was performed using the NovaSeq Illumina platform with a 150PE strategy at Novogene (Cambridge).

Data Processing

The GenErode pipeline v0.4.2 (Kutschera *et al*., 2022), suitable for processing both modern and historical short-read data, was used to trim and process all modern and historical sequence data along with highly contiguous reference genomes for *A. crataegi* (Ebdon *et al*., 2022) or *P. icarus* (Lohse, 2023). Both these reference genome assemblies are high quality, produced from single male specimens, with an assembled Z chromosome (lepidopteran males are commonly the homogametic sex). The *A. crataegi* assembly is ~ 230 Megabases (Mb) long and is scaffolded into 26 pseudochromosomes and a Z chromosome. The *P. icarus* assembly is ~512 Mb scaffolded into 23 pseudochromosomes and a Z.

FastQC v0.11.9 (Andrews *et al*., 2012) and fastp v0.20.1 (Chen *et al*., 2018) were used to trim adapters and merge paired sequences (for historical samples). Trimmed reads were aligned to reference genomes using BWA aln for hDNA and BWA mem for modern data (Li and Durbin, 2009) with default settings in the Generode Pipeline. Post-mapping, any samples sequenced across multiple lanes, were merged and PCR duplicates removed. MapDamage v2.0 (Jónsson *et al*., 2013) was used to estimate damage patterns in historical samples and rescale base quality scores. Finally, GATK IndelRealinger, (Van der Auwera and O'Connor, 2020) was used to remap any reads around indels to improve mapping accuracy. For the modern, *P. icarus* data, all resulting BAM files were subsampled to 10X to mitigate the influence of varying coverage on heterozygous calls. These BAM files were processed further to call variants and reconstructing historical demography PSMC (Li and Durbin, 2011). For the latter, we used full coverage data for the modern *P. icarus* samples.

Variants calls were made on a per-sample basis using Bcftools mpileup (Danecek *et al*., 2021) with a minimum depth threshold of 1/3 of average coverage (with an absolute minimum of 3X) and a maximum of ten times the average coverage while minimum mapping quality and minimum base quality were both set to 30 (-Q 30 -q 30). Additional filters as a part of the GenErode pipeline, included an allelic imbalance filter for heterozygous sites, removal of variants in repetitive regions of the genome, and removal of variants in CpG regions. A final VCF was generated that consisted of only biallelic single nucleotide polymorphisms (SNPS) and genotypes with no more than 10% missing data across all samples (historical *A. crataegi* samples, containing 305,394 SNPs) or no missing data (modern *P .icarus* data, containing 5,656,704 SNPs). SNPs from these VCFs were used for downstream analyses, unless otherwise stated. Our configuration files for the GenErode pipeline are available at https://github.com/rmwhitla/BVWpaper.

Population Structure

Population structure of both the historical *A. crataegi* and modern *P. icarus* samples were compared using principal component analysis (PCA) based on filtered genotypes in VCF format produced from the GenErode pipeline. Before PCA, any SNPs on the Z chromosome were filtered out. Furthermore, we pruned the remaining SNP sets for linkage disequilibrium over windows of 10 kilobases with a step size of 10 bases and an r^2 of 0.1 using Plink 1.9 (Purcell *et al.,* 2007) resulting in a total of 14,930 (*A. crataegi*) or 197,182 *(P. icarus*) SNPs. PCA was performed using Plink v1.9 with a minor allele frequency (MAF) threshold of 10% (historical *A. crataegi)* or 15% (modern *P. icarus*).

Given the low and variable coverage (2-13X) of the historical *A*. *crataegi* samples, we also performed PCA based on genotype likelihoods, which can account for statistical uncertainty in called genotypes arising from low coverage data (Korneliussen *et al*., 2014). Genotype likelihoods were calculated from BAM files produced as part of the GenErode pipeline, using ANGSD (Korneliussen *et al*., 2014) with a minimum mapping quality at 30 and minimum base quality at 20 (-minMapQ 30 -minQ 20) filtering only for the same sites used with the PCA for hard genotype calls. Pcangsd (Meisner and

Albrechtsen, 2018) was used for the PCA of genotype likelihoods with a minimum allele frequency of 10% (--minMaf 0.1).

Historical Demography Reconstruction

We used the Pairwise Sequentially Markovian Coalescent model as implemented in PSMC v0.6.5 (Li and Durbin, 2011) to determine if British populations of *A. crataegi* and *P. icarus* exhibit similar trends of historical change in effective population size (N_e), as would be expected from colonization of the British Isles after the last glacial maximum (LGM) *c.* 13,000-10,000 years ago (Hewitt, 1999). As input for PSMC, we generated a consensus fastq file from BAM files, using only samples with average coverage greater than 5X. Consensus fastq files were generated with Bcftools pileup and vcfutils.pl, filtering positions with minimum depth less than 5X or maximum depth greater than 100X in *A. crataegi* and with minimum depth less than 5X or maximum depth greater than 500X in *P. icarus*. Additionally, any positions in repeat masked regions or CpG regions (historical samples only) or Z chromosomes were also removed. PSMC was run with: number of iterations -N = 25, maximum coalescent time -t =15 and the atomic time interval (-p) at 66 (for *A. crataegi*: 4+25*2+3+9) and 66 (for *P. icarus*: 27*2+4+8). The atomic time interval was chosen to avoid overfitting by ensuring that at least 10 recombination events occurred over each interval after the 20th iteration of PSMC (Li and Durbin, 2011). To account for low coverage and stochastic loss of heterozygosity, the initial theta/rho ratio parameter for each sample was calculated based on false negative rates (FNR) corresponding to the coverage of that sample. FNR for different coverages were determined based on the recommendation of Li and Durbin (2011) (https://github.com/lh3/psmc), and as implemented in Sarabia *et al*. (2020). To obtain FNR estimated for *A. crataegi*, we used a subsampled version of a publicly available high coverage (23X) modern *A. crataegi* (accession: SRR7948941) specimen from Japan that was sequentially subsampled to 5X, 8X, 10X, 15X and 20X. For FNR correction for low coverage *P. icarus* samples, we used the same subsampling strategy as for *A. crataegi* using one of our high coverage samples (DGCM100, coverage: >50X). The initial theta/rho (-r) parameter for PSMC was then set to 5/(1-FNR). To visualise the variability in our PSMC estimates, we ran 100 bootstraps for each sample over 5 Mb segments using the same initial parameters as the original run for each sample.

To scale population size and times resulting from the PSMC analysis, we used a direct estimate of spontaneous mutation rate of 2.9 x10-^{9} from *Heliconius Melpomene* (Keightley *et al.,* 2015). For *A. crataegi*, which is univoltine in the geographic range our samples came from (Solonkin *et al.,* 2021), we used 1 generation per year. For *P. icarus*, we used 0.5 generations (two broods) per year, although the number of broods are variable in this species (one in Scotland but two or three in Southern England and parts of Europe). However, we tested the sensitivity of the PSMC curves for *P. icarus* by changing the generation time to 1 and 3 broods per year.

Genetic Diversity and Inbreeding

We used genome-wide heterozygosity is our measure of genetic diversity. We estimated genomic heterozygosity using ANGSD, which utilises a genotype likelihood approach that is more suitable for low coverage data (Korneliussen *et al*., 2014). We calculated genotype likelihoods as described above (PCA using genotype likelihoods), except in this case no LD pruning was performed. To estimate global heterozygosity, we calculated the unfolded site frequency spectrum (SFS) using the realSFS function in ANGSD.

Next, we estimated individual inbreeding coefficients based on the detection of the number and size of Runs of Homozygosity (RoH). RoH were calculated using Plink v1.9 (Purcell *et al.,* 2007). We used a sliding window size of 100 SNPs (homozyg-window-snp 100). The maximum number of heterozygotes per window was 2 (homozyg-window-het 2). Minimum SNP count for each RoH was 25 (homozyg-snp 25) and the minimum size was set to 50kb (homozyg-kb 50). We allowed 10 missing sites per window (intermediate filtering - homozyg-window-missing 10). For the maximum number of heterozygote sites within a RoH, we allowed 3 (homozyg-het3). Using these results, F_{RoH}, the inbreeding coefficient, was determined using the overall proportion of the genome in RoH. Results from Plink were summarised in R v.4.3.0 (R Core Team, 2023) using the package DetectRuns (Biscarini *et al.,* 2018). F_{ROH} was estimated for RoH ≥ 100 Kilobases (Kb), RoH ≥ 500 Kb or RoH ≥ 1 Megabase (Mb) for samples with average coverage ≥ 5X. R0H ≥ 100 Kb and less than 1 Mb may

result from background inbreeding due to drift whereas R0H $\geq$ 1 Mb result from recent consanguineous mating (Ceballos *et al*., 2018).

Genetic Load Estimation

To estimate the number of deleterious mutations or the genetic load we used snpEff (Cingolani *et al.,* 2012) to annotate genetic variants of high impact (likely highly disruptive to protein, e.g., loss of function (LoF) variants), moderate impact (e.g. missense variants) and low impact (synonymous, likely harmless but may be weakly deleterious under strong codon bias), as well as modifier (non-coding) variants. We focused on annotated variants within coding regions (LoF, missense or synonymous). snpEFF only predicts the potential phenotypic impact of the variants, and we lack any prior knowledge of the selection or dominance coefficients for these variants. Nonetheless, variants classified as LoF, missense or synonymous can be considered proxies for strongly, mildly or weakly deleterious variants (Dussex *et al*., 2023).

For the analysis of genetic load, we only used samples with $\geq$ 5X average depth and no missing data across samples. This resulted in 179,135 SNPs for historical *A. crataegi* and 5,656,704 SNPs for *P. icarus*. Gene annotations (as GTF files) for both species were retrieved from Ensembl's Darwin Tree of Life data portal (https://projects.ensembl.org/darwin-tree-of-life/). Both annotations were generated by the Ensembl Gene Annotation System (Aken *et al.*, 2017) optimised for non-vertebrates. The annotation for *A.crataegi* (Ebdon *et al.,* 2022) is based on a single male from Catalunya, Spain and the annotation for *P. icarus* (Lohse, 2023) is based on a single male from Scotland.

We also determined the relative excess in the frequency of shared derived alleles (ignoring the impact of *de novo* mutations) across the LoF, missense and synonymous categories in both species. We used an approach previously described by Xue *et al*. (2015) and implemented by van der Valk *et al*, (2019). For each variant category we calculate a ratio R_{xy} that is a measure of shared derived alleles with a higher frequency in population X (British) compared to population Y (European). An $R_{xy} > 1$ implies a relative increase in frequency in population and X compared to Y, whereas < 1 implies a

relative decrease in X relative to Y. To estimate confidence intervals for our R_{xy} ratios we used a jackknife approach dropping a single chromosome at a time.

Statistical Analysis

For all statistical analysis between historical specimens of *A. crataegi* we used a linear model with average coverage as a covariate. Marginal means of all estimates for each group were estimated using the R package emmeans (Lenth, 2023). For *P. icarus*, statistical comparisons were made using Welch's two sample t-tests. All statistical comparisons were performed in R v.4.3.0 (R Core Team, 2023).

Results

Resequencing of hDNA

The final sequencing depth of the 17 historical *A. crataegi* specimens varied between 2X-13X, with higher mean coverage of European samples (10.13X) compared to British specimens (6.03X) and sequencing depth varied from 2X-13X (Table S1). The proportion of the genome sequenced to a depth of 5X or greater was 50.77% (Table S1)

Postmortem degradation is seen in the 3' ends of our historical reads (Fig. S1) as in increase in G>A transversions. The 5' ends of our fragments show little to no sign of historical damage due to the application of the USER enzyme during library preparation. Mean fragment length of sequencing reads, after trimming and merging reads, was 49bp with average fragment length per sample ranging from 44-65bp (Table S1). There was a significant positive correlation between mean fragment length and age of sample (Fig. S2) with fragment lengths increasing with time (Mullin *et al.,* 2022).

In our extraction blanks (BL1 and BL2), 0.49% and 0.43% of reads mapped to the *A. crataegi* genome respectively. In the blanks, 0.2% of bases had 1X coverage, and 0% of bases had higher coverage.

The average read lengths in the two blanks were 91 and 97. Overall, analysis of the mapping profile of the blanks suggests no evidence for cross contamination during DNA extraction and library preparation of the historical samples.

Spatial Genetic Structuring between Great British and European Mainland Butterflies

We use a principal component analysis (PCA) of LD-pruned SNPs to examine the population structure of both *A. crataegi* and *P. icarus*. A PCA of all historical *A. crataegi* specimens suggests spatial genetic structuring across historical samples regardless of collection date (Fig. 1B). Given the low to moderate coverage of our historical specimens (Table S1), we also performed a PCA based on genotype likelihoods. The PCA based on genotype likelihoods (Fig. S3) yielded visually similar results to the one based on hard genotype calls (Fig. 1B).

PC1 explained 10.8% of the variation, while PC2 explained 8.6% of variation. Spatial genetic structuring is apparent along PC1 with European samples tightly clustering on one end with negative values for PC1, whereas points with greater values of PC1 represent specimens from range margins (Wales; see Fig. S3 for a labelled PCA- but note the axes are reversed) but also include a sample with unknown origins (K05L) and one specimen from Kent (NHM026). In contrast to the geographical cline apparent along PC1, there is no clear cline present along PC2 but the points with the most extreme values represent the most recent specimens from Kent (NHM928 and NHM019). *P. icarus* exhibit similar structuring along PC1 with the southern England sample (Fig. 1C) sorting in between the European and marginal Scottish samples. Overall, sorting of specimens along PC1 for both species reveals a tight clustering among European butterflies and more geographic structuring within GB samples.

Three historical *A. crataegi* specimens (NHM898, NHM027 and NHM918; Fig. 1B) labelled as GB clustered closer to French and Belgium samples rather than the GB samples. Rather than a case of mislabelling, these specimens may potentially be cases of European *A. crataegi* released in Britain - a not uncommon practise from the late 19th to early 20th century (Pratt, 1983) – or potential migrants.

For subsequent analyses, these three samples are grouped with the other mainland European samples.

Demographic Reconstruction

We used Pairwise Sequentially Markovian Coalescent models to reconstruct the historical demography of both species to compare consistency within populations and obtain approximations of the N_e of British populations at the time of post-glacial colonisation. Historical demographic reconstruction of *A. crataegi* using PSMC shows similar trajectories of effective population size (N_e; Fig. 2A, Fig. S4) for both British and European samples up to the last glacial maximum (LGM, *c.* 22 to 17 Kya). British *A. crataegi,* show a marked decrease in N_e (consistent with a potential bottleneck) around the LGM and Younger Dryas (12.8–11.5 Kya). A similar pattern of trajectories is also observed in *P. icarus* (Fig. 2B, Fig. S5) but, the divergence in N_e appears to happen earlier, just before the LGM, but this trend is sensitive to the number of generations per year (Fig. S6). Both reconstructions indicate British *A. crataegi* and *P. icarus* show reduced N_e before or around 12 Kya.

Approximation of N_e at the time of colonisation are ~5 x 10^4 for *A. crataegi,* and ~2 x 10^5 for *P. icarus* (Fig. 2). We note that although we used higher coverage samples to estimate the FNR for low coverage samples, for samples below 8X (*A. crataegi*) or 14X (*P. icarus*) we could not recover N_e estimated for the high coverage data before 50 kya (Fig. S7 and S8). However, this discrepancy was limited to the magnitude of N_e and not the trajectories themselves.

Estimates of heterozygosity and inbreeding

Individual genome-wide heterozygosity (Fig. 3A-B) rates were lower for British samples for both *A. crataegi* and *P. icarus*. Heterozygosity between British and European *A. crataegi* was significantly different ($F_{1,14}$=25.21, p=0.000187, Fig. 3A), after accounting for variability in sample coverage. British *A. crataegi* exhibited an average reduction of 17.9% in heterozygosity relative to European samples. Heterozygosity was also significantly different between the British and French *P.icarus*

(t=9.26, *degrees of freedom* (df) = 3.52, p= 0.001356; Fig. 3B) samples, with an average reduction of 14.5%.

For detecting RoHs we used three thresholds: ≥ 100 kb, ≥ 500 kb or ≥ 1 Mb. British *A. crataegi* showed a much higher proportion of the genome in RoH (17% in ≥ 100 Kb, 3.7% in ≥ 500 kb and 2.2% in ≥ 1 Mb; Fig. 3C) than European *A. crataegi* (1% in ≥ 100 Kb, 0.2% in ≥ 500 kb, and 0.1% in ≥ 1 Mb; Fig. 3C). Additionally, the difference in all RoH categories were significantly different (≥ 100 Kb: $F_{1,10}$=36.53, p=0.0001246; ≥ 500 kb: $F_{1,10}$=28.47 p=0.00033; ≥ 1MB: $F_{1,10}$=6.91, p=0.025) with all six British *A. crataegi* harbouring at least one RoH ≥ 1 Mb (min F_{RoH} = 0.8%, max F_{RoH} = 5.7%) , whereas only one of seven amongst European specimens harboured an RoH ≥ 1 Mb . In contrast, British *P. icarus* had a relatively smaller proportion (3%) of its genome in RoHs ≥ 100 Kb, 0.2% in ≥ 500 Kb, and none in regions ≥ 1 Mb (Fig. 3D), while French *P. icarus* had negligible RoHs in ≥ 100 Kb (0.02%) and none in ≥ 500 kb or ≥ 1 Mb. The difference in in RoH segments ≥ 100 Kb between British and French *P. icarus* was marginally significant (t=-3.15, df = 3, p= 0.0507; Fig. 3D).

Estimation of genetic load

To estimate genetic load, we used comparisons of total numbers of LoF, missense and synonymous variants between British and European specimens. In *A. crataegi* (Fig. 4A), the mean number of LoF variants was not significantly different ($F_{1,10}$=2.8, p=0.125) between the British (17.5) and European specimens (20.2). However, the total numbers of missense and synonymous were significantly lower in British *A. crataegi (*Fig. 4B & C*).* In *P. icarus,* all variant classes were significantly lower in British samples (Fig. 4D-C). This suggests either no change (LoF) or reduction (missense and synonymous) in genetic load in British specimens.

Next, we sought to determine if there was any evidence for increase in realised load (deleterious variants exposed in homozygosity, assuming most deleterious variants are partially recessive) in British specimens. We did so by comparing the predicted variant classes in homozygous (Fig. S9A-C) state. For *A. crataegi* there was no significant increase homozygosity for LoF and missense classes, but there was a statistically significant increase in synonymous variants ($F_{1,10}$ = 6.37, p= 0.03)

between the British specimens (893) and European (864) specimens. British *P. icarus* show no significant change in homozygosity across all three variant classes (Fig S10).

Finally, we analysed genetic load in terms of relative increase in derived alleles. For *A. crataegi* we observed an overall decrease in the number of LoF variants (Fig. 5A) in British specimens but a statistically significant increase in frequencies, relative to continental specimens, in missense and synonymous variants. Although, the difference in LoF variants is not significantly different from 1 (equal frequency between populations), there is a trend of decrease (Fig. 5A). Shared derived alleles in British *P. icarus* (Fig. 5B) exhibit a significant decrease in frequency relative to continental specimens, however, the change in LoF frequency was not significantly different from 1.

Discussion

We used historical genomes of *A. crataegi* to provide a genomic perspective on their extirpation from the British Isles in the early twentieth century. We complemented our analysis with a similar approach using *P. icarus*, a butterfly with long-term stable population size within the British Isles. Our results provide genomic insights into the extirpation of *A. crataegi* and the utility of genomic erosion metrics to identify vulnerable insect populations. This approach should be applicable to other insect species using both museum and contemporary specimens. Comparative data on genomic erosion in threatened, stable and expanding species, but also those with different life histories could improve the predictive capacity for using metrics of genome erosion for identifying and managing threatened populations and species (Kryiazis et al., 2023; Webster et al., 2023).

British butterflies are genetically distinct from continental European populations and share a common demographic history

Our population structure analysis (Fig. 1B-C) and PSMC-based demographic reconstruction (Fig. 2) suggest that British populations of both *A. crataegi* and *P. icarus* are genetically distinct from their continental European counterparts. However, there was evidence of strong structuring within the British populations (Fig. 1B-C). Structuring apparent within the British specimens could result from

recent gene flow with European mainland populations or alternatively, from sequential founder events and allele surfing after colonisation and subsequent expansion (Excoffier and Ray, 2008) or a combination of these two hypotheses. Gene flow may be a more likely explanation for the structuring observed in *P. icarus* as has been suggested by Arif *et al*., (2021). However, for *A. crataegi*, the consistency of demographic reconstructions (Fig. 2 and Fig S4), genome-wide heterozygosity (Fig. 2A), RoHs (Fig. 2C) and genetic load (Fig. 3A-C) amongst the British specimens, do not favour a hypothesis of gene flow or admixture from the continent.

The post-glacial colonisation of the British Isles would likely have commenced around ~12 Kya following the Younger Dryas (Hewitt, 1999). The colonisation and expansion of the British Isles likely resulted in a genetic bottleneck due to a small founding population. Using allozymes, bottlenecks associated with post-glacial colonization events have been previously suggested for *P. icarus* (DeKeyser *et al*., 2012). Our demographic reconstructions suggest that both species, *A. crataegi* and *P. icarus* (Fig. 2, Fig. S4, Fig. S5), underwent a decrease in N_e around the passing of the Eemian interglacial period (~125 Kya) until around the last glacial maximum (LGM, ~20 Kya). After this time (~15 Kya) British *A. crataegi* exhibit a marked reduction in N_e, consistent with a genetic bottleneck associated with a founder event. At the same time, European *A. crataegi* show a drastic increase in N_e presumably because of mixing of distinct populations previously isolated in glacial refugia (Hewitt, 1999; Hinojosa *et al.,* 2019).

P. icarus specimens also exhibit a visually similar pattern of N_e trajectories as *A. crataegi* (but not magnitude of N_e) between British and European specimens. However, in this case the divergence between British and European specimens appear to occur before the LGM (Fig. 2B, Fig. S4, Fig. S5). This discrepancy in timing of shifts in trajectories may result from variability in generation time for *P. icarus* (Fig. S6) which can vary from 1 to 3 broods a year.

These results suggest that both species underwent a bottleneck during the post-glacial colonisation of the British Isles. However, the founding N_e for *A. crataegi* is ~4X smaller than *P. icarus* (Fig. 2), on the order of $\approx 10^4$ which is a magnitude smaller than the commonly observed sizes of 10^5 or greater reported in butterflies (Mackintosh *et al.*, 2019). It is possible that lack of recovery from a small

founding N_e contributed to *A. crataegi's* eventual demise in Britain, however differences in thermal tolerances or other unknown ecological traits between the *A. crataegi* and *P. icarus* also potentially influenced their contrasting fates. Additionally, PMSC estimates of N_e are not reliable within the last 10 Kya but levels on inbreeding can be more informative of contemporary N_e.

Genomic symptoms of the demise of* A. crataegi *in Britain

The last recorded sighting of *A. crataegi* in Britain dates from *c.* 1925 (Pratt, 1983). However, by this time, the species was thought to be limited to the south-eastern regions of Kent and had already disappeared from other parts of its range as early as 1880 (Pratt, 1983 and references therein). Consistent with these observations, three of 13 of our British samples, samples collected in 1908, 1924 and 1888, appear to be of European origin in terms of population structure, demographic history, and genomic erosion. These specimens may have been captive-bred or wild-caught European (potentially from imported German larvae (Pratt, 1983)) stock released into British habitats or potential migrants.

There was a reduction in genome-wide heterozygosity in British specimens of both species, but reductions (compared to European populations) in both the extirpated *A. crataegi* and the extant and stable *P. icarus* were roughly comparable (17.9 and 14.5%, respectively). As a comparison, the observed reduction in genome-wide heterozygosity in the extinct *Glaucopsyche xerces* (Xerces Blue butterfly; de-Dios *et al*., 2023) was 22%, hence greater than what we observed for British *A. crataegi*. However, the former was calculated in comparison to a sister species (*Glaucopsyche lygdamus*) and hence it is unclear whether these estimates of reductions in genome-wide heterozygosity are comparable.

More conspicuously, we saw significant increases in homozygous regions of small (≥ 100 Kb), intermediate (≥ 500 Kb) and large (≥ 1 Mb) variety (Fig. 2) in British *A. crataegi*. The presence of small RoHs of would be consistent with ancestral inbreeding as also suggested by the small N_e estimated using PSMC. However, the presence of intermediate and large RoHs raise the possibility of more recent inbreeding (Fig. 2A), potentially as recent as 10s of generations ago (indicated by presence of

RoH ≥ 1 Mb). This distribution of RoHs would be consistent with spasmodic decline of *A. crataegi* in Britain. Evidence of more recent inbreeding supports a further, more recent bottleneck event, for e.g., pathogen infection, extreme seasonal conditions (Pratt 1983), that may have led to the species eventual demise in the British Isles. In contrast, no category of F_{RoH} was significantly different between British and continental European *P. icarus* suggesting this population likely recovered from the initial bottleneck at the time of post-glacial colonisation of Britain. Morever, F_{roh} for British *A. crataegi* are substantially larger in size and proportion than those estimated from historical genomes of the extinct *G. xerces* (de-Dios *et al*., 2023).

Our analyses of genetic load also suggest an increased number of shared derived missense and synonymous deleterious variants but not LoF variants (Fig. 5) in British *A. crataegi*. Accumulation of mildly or weakly deleterious variants, solely due to drift after a bottleneck, is expected to lead to decrease in the total number of such variants but increase in homozygosity across all variant classes (Dussex *et al.,* 2023). However, purging of deleterious variants through purifying selection should lead to a decrease in strongly deleterious variants (LoF) in both homozygous and heterozygous state, whereas homozygosity of mildly (missense) or weakly (synonymous) can still increase and inflate the genetic load. The accumulation of genetic load in small populations has been observed in many vertebrate populations (see references in Dussex *et al*. 2023) and in the Xerces blue butterfly (de-Dios *et al*., 2023). However, there has also been evidence documentation of purging of the most deleterious variants and accumulation of less deleterious variants in the same system in Alpine Ibex (Grossen *et al*., 2020) and Indian tigers (Khan *et al*., 2021) and is supported by simulations (Dussex *et al*., 2023). Although we observed no clear evidence of purging of LoF in British *A. crataegi* (Fig. 4A, Fig. S9A and D, Fig. 5A), we did see an increase in shared derived allele frequency in the missense and synonymous categories which may have contributed to extirpation by reducing fitness and population viability. Additionally, we also observed statistically significant increase in realised load but this was limited to only synonymous variants. Although synonymous variants are normally classified as neutral, they can be weakly deleterious under codon bias and Lepidoptera do exhibit an A/T codon usage bias (Näsvall *et al*., 2023). In contrast, *P. icarus* displayed a decrease in realised load (Fig. S10) and a decrease in frequencies of shared derived alleles of the missense and synonymous categories (Fig. 5B), suggesting these populations may have been less influenced by genetic drift.

These results raise possibility that the accumulation of mildly and weakly deleterious mutations contributed to the decline of *A. crataegi* in Britain by triggering a mutational meltdown (Lynch et al., 1995b). However, it should be noted that the actual fitness effect including selection and dominance coefficients of the annotated variants remains unknown. Moreover, given our limited sample sizes and lack of statistical significance in some cases, our results on genetic load should be interpreted with caution. Finally, given that the annotation for *P. icarus* (Lohse, 2023) is based on a single male from Scotland, our predictions of phenotypic impact of variants for this species may suffer from annotation bias, for example, by reducing the number of predicted LoF and missense variants.

Overall, we provide compelling evidence that *A. crataegi* suffered a severe bottleneck during the colonisation of the British Isles. The distribution of RoHs in this population suggest that N_e likely remained low following the bottleneck, potentially through more bottleneck events in the 19th century, with the accumulation of weak or mild deleterious mutations and potential increase in realised load, which may have further exacerbated population persistence in the short term. Interestingly, two other species, *Lycaena dispar* (Large copper butterfly) and *Cyaniris semiargus* (Mazarine blue butterfly) were also extirpated in Britain in the mid to late 19th century. It would be instructive to examine whether genomes of these extirpated population exhibit similar symptoms of genomic erosion observed in *A. crataegi*.

Outlook

Conservation genomics could be an extremely helpful tool in identifying populations or species at risk of extinction or extirpation. This could be particularly relevant for invertebrates, whereas opposed to vertebrates, reductions in census population size may not be easily monitored (Cardoso *et al*., 2011; Cardoso and Leather, 2019; Didham *et al.,* 2020). Genetic diversity is one component of genomic erosion and important for species persistence and adaptation and is currently recognized as one of the three relevant levels of biological diversity by the Convention on Biological Diversity (CBD, O'Brien *et al*., 2022). Genetic diversity is relatively simple to determine from a small number of samples with a handful of genetic markers (Bosse and Loon, 2022). However, genomic heterogeneity in genetic diversity (Robinson *et al.,* 2019) and its functional relevance to species persistence and

adaptation (Teixeira and Huber, 2021) are from clear. Hence, a more meaningful approach to identifying vulnerable species would require genome-wide estimates of inbreeding and genetic load, measures that have more direct links to population fitness. We, as others (e.g., Hohenlohe *et al*., 2021), recommend that single metrics such as genetic diversity (individual heterozygosity in our case) may not be sufficient on their own to delineate threatened populations and need to be supplemented with estimates of inbreeding and genetic load. Here and elsewhere (de-Dios *et al*., 2023), we demonstrate that this achievable for insect species with a relatively small number of samples.

To evaluate how genomic erosion translates into population persistence or extinction we require further empirical studies of not only declining but stable and expanding populations as well. In this respect, isolated populations on islands, including the British Isles, offer an excellent system to study the genomics of populations with contrasting fates. Museum collections can provide a substantial and unequivocal source of threatened populations (e.g., samples from extirpated or extinct populations). Recent advances in resequencing of museum samples (Korlević *et al*., 2021, Mullin *et al*., 2023), bioinformatic tools (e.g., Korneliussen *et al*., 2014, Kutschera *et al.,* 2022) and the availability of high-quality genomes with annotations (e.g., Earth Biogenome Project, Darwin Tree of Life Project) are providing much needed resources for such work. Although challenges remain, for example, evaluating and interpreting fitness effects of deleterious mutations (de Valles-Ibáñez *et al*., 2016; Robinson *et al*., 2023), we expect museomics in combination with contemporary samples will expand our understanding of the genomics of extinction but also improve capacity to monitor and manage threatened insect populations (Díez-del-Molino *et al.,* 2018; Jensen *et al.*, 2022).

Acknowledgements

This work was supported by a Nigel Groome PhD Studentship (Department of Biological and Medical Sciences, Oxford Brookes University) to RW and a pilot funding grant from the Centre for Functional Genomics, Oxford Brookes University to SA. We would like to thank James Hogan at the Oxford University Natural History Museum (OUNHM) and Geoff Martin at the Natural History Museum, London (NHM) for their help in obtaining the Black-veined white museum samples. We would also like to thank

Joe Middleton-Welling for obtaining the samples for a pilot study from the OUNHM and Konrad Lohse for providing genomic DNA and the Darwin Tree of Life for the Black-veined White genome assembly.

Author Contributions

TGS, SA and RW designed the study. TGS, SA and RW collected data. RW completed all molecular work with supervision and support from KH, CB and SA. RW analysed all the data with the help of SA. SA and RW wrote the paper with input from all authors.

Data Availability Statement

Raw sequence data is available under project accessing PRJEB70473 (for *A. crataegi*) and PRJEB70726 (for *P. icarus*) on the European Nucleotide Archive (ENA). Configuration files for GenErode and additional script files are available at https://github.com/rmwhitla/BVWpaper. Protocol for library preparation of historical samples is available at https://www.protocols.io/view/library-prep-for-cut-run-with-nebnext-ultra-ii-dna-bwgjpbun)

References

Aken, B.L., Achuthan, P., Akanni, W., Amode, M.R., Bernsdorff, F., Bhai, J., Billis, K., Carvalho-Silva, D., Cummins, C., Clapham, P., Gil, L., Girón, C.G., Gordon, L., Hourlier, T., Hunt, S.E., Janacek, S.H., Juettemann, T., Keenan, S., Laird, M.R., Lavidas, I., Maurel, T., McLaren, W., Moore, B., Murphy, D.N., Nag, R., Newman, V., Nuhn, M., Ong, C.K., Parker, A., Patricio, M., Riat, H.S., Sheppard, D., Sparrow, H., Taylor, K., Thormann, A., Vullo, A., Walts, B., Wilder, S.P., Zadissa, A., Kostadima, M., Martin, F.J., Muffato, M., Perry, E., Ruffier, M., Staines, D.M., Trevanion, S.J., Cunningham, F., Yates, A., Zerbino, D.R., Flicek, P., 2017. Ensembl 2017. Nucleic Acids Research 45, D635–D642. https://doi.org/10.1093/nar/gkw1104

Andrews, S., 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Arif, S., Gerth, M., Hone-Millard, W.G., Nunes, M.D.S., Dapporto, L., Shreeve, T.G., 2021. Evidence for multiple colonisations and Wolbachia infections shaping the genetic structure of the widespread butterfly Polyommatus icarus in the British Isles. Molecular Ecology 30, 5196–5213. https://doi.org/10.1111/mec.16126

Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O.U., Swartz, B., Quental, T.B., Marshall, C., McGuire, J.L., Lindsey, E.L., Maguire, K.C., Mersey, B., Ferrer, E.A., 2011. Has the Earth's sixth mass extinction already arrived? Nature 471, 51–57. https://doi.org/10.1038/nature09678

Biscarini, F., Cozzi, P., Gaspa, G., Marras, G., 2018. detectRUNS: Detect runs of homozygosity and runs of heterozygosity in diploid genomes [WWW Document]. URL https://cran.r-project.org/web/packages/detectRUNS/index.html (accessed 11.2.23).

Bortoluzzi, C., Wright, C.J., Lee, S., Cousins, T., Genez, T.A.L., Martin, F.J., Haggerty, L., Durbin, R., n.d. Lepidoptera genomics based on 88 chromosomal reference sequences informs population genetic parameters for conservation.

Bosse, M., van Loon, S., 2022. Challenges in quantifying genome erosion for conservation. Front Genet 13, 960958. https://doi.org/10.3389/fgene.2022.960958

Brooks, A.J., Bradley, S.L., Edwards, R.J., Goodwyn, N., 2011. The palaeogeography of Northwest Europe during the last 20,000 years. Journal of Maps 7, 573–587. https://doi.org/10.4113/jom.2011.1160

Butchart, S., Butchart, S., Stattersfield, A.J., Brooks, T., 2006. Going or gone: defining "Possibly Extinct" species to give a truer picture of recent extinctions. Bulletin of the British Ornithologists' Club 126A, 7--24.

Cardoso, P., Erwin, T.L., Borges, P.A.V., New, T.R., 2011. The seven impediments in invertebrate conservation and how to overcome them. Biological Conservation 144, 2647–2655. https://doi.org/10.1016/j.biocon.2011.07.024

Cardoso, P., Leather, S.R., 2019. Predicting a global insect apocalypse. Insect Conservation and Diversity 12, 263–267. https://doi.org/10.1111/icad.12367

Caughley, G., 1994. Directions in Conservation Biology. Journal of Animal Ecology 63, 215–244. https://doi.org/10.2307/5542

Ceballos, G., Ehrlich, P.R., Barnosky, A.D., García, A., Pringle, R.M., Palmer, T.M., 2015. Accelerated modern human-induced species losses: Entering the sixth mass extinction. Sci Adv 1, e1400253. https://doi.org/10.1126/sciadv.1400253

Ceballos, F.C., Joshi, P.K., Clark, D.W., Ramsay, M., Wilson, J.F., 2018. Runs of homozygosity: windows into population history and trait architecture. Nat Rev Genet 19, 220–234. https://doi.org/10.1038/nrg.2017.109

Charlesworth, D., Willis, J.H., 2009. The genetics of inbreeding depression. Nat Rev Genet 10, 783–796. https://doi.org/10.1038/nrg2664

Chen, S., Zhou, Y., Chen, Y., Gu, J., 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34, i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., Ruden, D.M., 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly (Austin) 6, 80–92. https://doi.org/10.4161/fly.19695

Cowie, R.H., Bouchet, P., Fontaine, B., 2022. The Sixth Mass Extinction: fact, fiction or speculation? Biological Reviews 97, 640–663. https://doi.org/10.1111/brv.12816

Crossley, M.S., Smith, O.M., Berry, L.L., Phillips-Cosio, R., Glassberg, J., Holman, K.M., Holmquest, J.G., Meier, A.R., Varriano, S.A., McClung, M.R., Moran, M.D., Snyder, W.E., 2021. Recent climate change is creating hotspots of butterfly increase and decline across North America. Global Change Biology 27, 2702–2714. https://doi.org/10.1111/gcb.15582

Dale, C.W., 1887. Query regarding *Aporia crataegi*. The Entomologist's Monthly Magazine 23, 214.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., Li, H., 2021. Twelve years of SAMtools and BCFtools. Gigascience 10, giab008. https://doi.org/10.1093/gigascience/giab008

Darwin Tree of Life Project Consortium, 2022. Sequence locally, think globally: The Darwin Tree of Life Project. Proc Natl Acad Sci U S A 119, e2115642118. https://doi.org/10.1073/pnas.2115642118

Davis, C.L., Guralnick, R.P., Zipkin, E.F., 2023. Challenges and opportunities for using natural history collections to estimate insect population trends. Journal of Animal Ecology 92, 237–249. https://doi.org/10.1111/1365-2656.13763

de-Dios, T., Fontsere, C., Renom, P., Stiller, J., Llovera, L., Uliano-Silva, M., Sánchez-Gracia, A., Wright, C., Lizano, E., Caballero, B., Navarro, A., Civit, S., Robbins, R.K., Blaxter, M., Marquès-Bonet, T., Vila, R., Lalueza-Fox, C., 2023. Whole-genomes from the extinct Xerces Blue butterfly can help identify declining insect species. eLife 12. https://doi.org/10.7554/eLife.87928

De Keyser, R., Shreeve, T.G., Breuker, C.J., Hails, R.S., Schmitt, T., 2012. Polyommatus icarus butterflies in the British Isles: evidence for a bottleneck. Biological Journal of the Linnean Society 107, 123–136. https://doi.org/10.1111/j.1095-8312.2012.01925.x

de Manuel, M., Barnett, R., Sandoval-Velasco, M., Yamaguchi, N., Garrett Vieira, F., Zepeda Mendoza, M.L., Liu, Shiping, Martin, M.D., Sinding, M.-H.S., Mak, S.S.T., Carøe, C., Liu, Shanlin, Guo, C., Zheng, J., Zazula, G., Baryshnikov, G., Eizirik, E., Koepfli, K.-P., Johnson, W.E., Antunes, A., Sicheritz-Ponten, T., Gopalakrishnan, S., Larson, G., Yang, H., O'Brien, S.J., Hansen, A.J., Zhang, G., Marques-Bonet, T., Gilbert, M.T.P., 2020. The evolutionary history of extinct and living lions. Proceedings of the National Academy of Sciences 117, 10927–10934. https://doi.org/10.1073/pnas.1919423117

de Valles-Ibáñez, G., Hernandez-Rodriguez, J., Prado-Martinez, J., Luisi, P., Marquès-Bonet, T., Casals, F., 2016. Genetic Load of Loss-of-Function Polymorphic Variants in Great Apes. Genome Biology and Evolution 8, 871–877. https://doi.org/10.1093/gbe/evw040

Dennis, R.L.H., Shreeve, T.G., Van Dyck, H., 2003. Towards a Functional Resource-Based Concept for Habitat: A Butterfly Biology Viewpoint. Oikos 102, 417–426.

Denton, G.H., Anderson, R.F., Toggweiler, J.R., Edwards, R.L., Schaefer, J.M., Putnam, A.E., 2010. The Last Glacial Termination. Science 328, 1652–1656. https://doi.org/10.1126/science.1184119

DeWoody, J.A., Harder, A.M., Mathur, S., Willoughby, J.R., 2021. The long-standing significance of genetic diversity in conservation. Molecular Ecology 30, 4147–4154. https://doi.org/10.1111/mec.16051

Díaz, S., Settele, J., Brondízio, E.S., Ngo, H.T., Agard, J., Arneth, A., Balvanera, P., Brauman, K.A., Butchart, S.H.M., Chan, K.M.A., Garibaldi, L.A., Ichii, K., Liu, J., Subramanian, S.M., Midgley, G.F., Miloslavich, P., Molnár, Z., Obura, D., Pfaff, A., Polasky, S., Purvis, A., Razzaque, J., Reyers, B., Chowdhury, R.R., Shin, Y.-J., Visseren-Hamakers, I., Willis, K.J., Zayas, C.N., 2019. Pervasive human-driven decline of life on Earth points to the need for transformative change. Science 366, eaax3100. https://doi.org/10.1126/science.aax3100

Didham, R.K., Basset, Y., Collins, C.M., Leather, S.R., Littlewood, N.A., Menz, M.H.M., Müller, J., Packer, L., Saunders, M.E., Schönrogge, K., Stewart, A.J.A., Yanoviak, S.P., Hassall, C., 2020. Interpreting insect declines: seven challenges and a way forward. Insect Conservation and Diversity 13, 103–114. https://doi.org/10.1111/icad.12408

Díez-del-Molino, D., Sánchez-Barreiro, F., Barnes, I., Gilbert, M.T.P., Dalén, L., 2018. Quantifying Temporal Genomic Erosion in Endangered Species. Trends in Ecology & Evolution 33, 176–185. https://doi.org/10.1016/j.tree.2017.12.002

Dincă, V., Montagud, S., Talavera, G., Hernández-Roldán, J., Munguira, M.L., García-Barros, E., Hebert, P.D.N., Vila, R., 2015. DNA barcode reference library for Iberian butterflies enables a continental-scale preview of potential cryptic diversity. Sci Rep 5, 12395. https://doi.org/10.1038/srep12395

Dussex, N., Morales, H.E., Grossen, C., Dalén, L., Oosterhout, C. van, 2023. Purging and accumulation of genetic load in conservation. Trends in Ecology & Evolution 38, 961–969. https://doi.org/10.1016/j.tree.2023.05.008

Dussex, N., van der Valk, T., Morales, H.E., Wheat, C.W., Díez-del-Molino, D., von Seth, J., Foster, Y., Kutschera, V.E., Guschanski, K., Rhie, A., Phillippy, A.M., Korlach, J., Howe, K., Chow, W., Pelan, S., Mendes Damas, J.D., Lewin, H.A., Hastie, A.R., Formenti, G., Fedrigo, O., Guhlin, J., Harrop, T.W.R., Le Lec, M.F., Dearden, P.K., Haggerty, L., Martin, F.J., Kodali, V., Thibaud-Nissen, F., Iorns, D., Knapp, M., Gemmell, N.J., Robertson, F., Moorhouse, R., Digby, A., Eason, D., Vercoe, D., Howard, J., Jarvis, E.D., Robertson, B.C., Dalén, L., 2021. Population genomics of the critically endangered kākāpō. Cell Genomics 1, 100002. https://doi.org/10.1016/j.xgen.2021.100002

Ebdon, S., Mackintosh, A., Lohse, K., Hayward, A., Arif, S., Whitla, R., 2022. The genome sequence of the black-veined white butterfly, Aporia crataegi (Linnaeus, 1758). https://doi.org/10.12688/wellcomeopenres.17709.1

Excoffier, L., Foll, M., Petit, R.J., 2009. Genetic Consequences of Range Expansions. Annual Review of Ecology, Evolution, and Systematics 40, 481–501. https://doi.org/10.1146/annurev.ecolsys.39.110707.173414

Excoffier, L., Ray, N., 2008. Surfing during population expansions promotes genetic revolutions and structuration. Trends in Ecology & Evolution 23, 347–351. https://doi.org/10.1016/j.tree.2008.04.004

Formenti, G., Theissinger, K., Fernandes, C., Bista, I., Bombarely, A., Bleidorn, C., Ciofi, C., Crottini, A., Godoy, J.A., Höglund, J., Malukiewicz, J., Mouton, A., Oomen, R.A., Paez, S., Palsbøll, P.J., Pampoulie, C., Ruiz-López, María J., Svardal, H., Theofanopoulou, C., de Vries, J., Waldvogel, A.-M., Zhang, Guojie, Mazzoni, C.J., Jarvis, E.D., Bálint, M., Formenti, G., Theissinger, K., Fernandes, C., Bista, I., Bombarely, A., Bleidorn, C., Čiampor, F., Ciofi, C., Crottini, A., Godoy, J.A., Hoglund, J., Malukiewicz, J., Mouton, A., Oomen, R.A., Paez, S., Palsbøll, P., Pampoulie, C., Ruiz-López, María José, Svardal, H., Theofanopoulou, C., de Vries, J., Waldvogel, A.-M., Zhang, Goujie, Mazzoni, C.J., Jarvis, E., Bálint, M., Aghayan, S.A., Alioto, T.S., Almudi, I., Alvarez, N., Alves, P.C., Amorim, I.R., Antunes, A., Arribas, P., Baldrian, P., Berg, P.R., Bertorelle, G., Böhne, A., Bonisoli-Alquati, A., Boštjančić, L.L., Boussau, B., Breton, C.M., Buzan, E., Campos, P.F., Carreras, C., Castro, L.Fi., Chueca, L.J., Conti, E., Cook-Deegan, R., Croll, D., Cunha, M.V., Delsuc, F., Dennis, A.B., Dimitrov, D., Faria, R., Favre, A., Fedrigo, O.D., Fernández, R., Ficetola, G.F., Flot, J.-F., Gabaldón, T., Galea Agius, D.R., Gallo, G.R., Giani, A.M., Gilbert, M.T.P., Grebenc, T., Guschanski, K., Guyot, R., Hausdorf, B., Hawlitschek, O., Heintzman, P.D., Heinze, B., Hiller, M., Husemann, M., Iannucci, A., Irisarri, I., Jakobsen, K.S., Jentoft, S., Klinga, P., Kloch, A., Kratochwil, C.F., Kusche, H., Layton, K.K.S., Leonard, J.A., Lerat, E., Liti, G., Manousaki, T., Marques-Bonet, T., Matos-Maraví, P., Matschiner, M., Maumus, F., Mc Cartney, A.M., Meiri, S., Melo-Ferreira, J.,

Mengual, X., Monaghan, M.T., Montagna, M., Mysłajek, R.W., Neiber, M.T., Nicolas, V., Novo, M., Ozretić, P., Palero, F., Pârvulescu, L., Pascual, M., Paulo, O.S., Pavlek, M., Pegueroles, C., Pellissier, L., Pesole, G., Primmer, C.R., Riesgo, A., Rüber, L., Rubolini, D., Salvi, D., Seehausen, O., Seidel, M., Secomandi, S., Studer, B., Theodoridis, S., Thines, M., Urban, L., Vasemägi, A., Vella, A., Vella, N., Vernes, S.C., Vernesi, C., Vieites, D.R., Waterhouse, R.M., Wheat, C.W., Wörheide, G., Wurm, Y., Zammit, G., 2022. The era of reference genomes in conservation genomics. Trends in Ecology & Evolution 37, 197–202. https://doi.org/10.1016/j.tree.2021.11.008

Fox, R., Dennis, E.B., Brown, A.F., Curson, J., 2022. A revised Red List of British butterflies. Insect Conserv Diversity 15, 485–495. https://doi.org/10.1111/icad.12582

Gomulkiewicz, R., Holt, R.D., 1995. When does Evolution by Natural Selection Prevent Extinction? Evolution 49, 201–207. https://doi.org/10.2307/2410305

Grossen, C., Guillaume, F., Keller, L.F., Croll, D., 2020. Purging of highly deleterious mutations through severe bottlenecks in Alpine ibex. Nat Commun 11, 1001. https://doi.org/10.1038/s41467-020-14803-1

Hayhow, D.B., Eaton, M.A., Stanbury, A.J., Burns, F., Kirby, W.B., Bailey, N., Beckmann, B., Bedford, J., Boersch-Supan, P.H., Coomber, F., Dennis, E.B., Dolman, S.J., Dunn, E., Hall, J., Harrower, C., Hatfield, J.H., Hawley, J., Haysom, K., Hughes, J., Johns, D.G., Mathews, F., McQuatters-Gollop, A., Noble, D.G., Outhwaite, C.L., Pearce-Higgins, J.W., Pescott, O.L., Powney, G.D., Symes, N., 2019. State of Nature 2019 [WWW Document]. URL https://nbn.org.uk/stateofnature2019/reports/ (accessed 10.25.23).

Hedrick, P.W., Lacy, R.C., Allendorf, F.W., Soule, M.E., 1996. Directions in Conservation Biology: Comments on Caughley. Conservation Biology 10, 1312–1320.

Hewitt, G.M., 1999. Post-glacial re-colonization of European biota. Biological Journal of the Linnean Society 68, 87–112. https://doi.org/10.1111/j.1095-8312.1999.tb01160.x

Hill, G.M., Kawahara, A.Y., Daniels, J.C., Bateman, C.C., Scheffers, B.R., 2021. Climate change effects on animal ecology: butterflies and moths as a case study. Biol Rev Camb Philos Soc 96, 2113–2126. https://doi.org/10.1111/brv.12746

Hinojosa, J.C., Koubínová, D., Szenteczki, M.A., Pitteloud, C., Dincă, V., Alvarez, N., Vila, R., 2019. A mirage of cryptic species: Genomics uncover striking mitonuclear discordance in the butterfly Thymelicus sylvestris. Molecular Ecology 28, 3857–3868.

Hohenlohe, P.A., Funk, W.C., Rajora, O.P., 2021. Population genomics for wildlife conservation and management. Mol Ecol 30, 62–82. https://doi.org/10.1111/mec.15720

IPBES, 2009. Summary for policymakers of the global assessment report on biodiversity and ecosystem services. IPBES Plenary at its seventh session (IPBES 7, Paris, 2019) , 29 April - 4 May 2019. https://doi.org/10.5281/zenodo.3553579

Jensen, E.L., Díez-del-Molino, D., Gilbert, M.T.P., Bertola, L.D., Borges, F., Cubric-Curik, V., de Navascués, M., Frandsen, P., Heuertz, M., Hvilsom, C., Jiménez-Mena, B., Miettinen, A., Moest, M., Pečnerová, P., Barnes, I., Vernesi, C., 2022. Ancient and historical DNA in conservation policy. Trends in Ecology & Evolution 37, 420–429. https://doi.org/10.1016/j.tree.2021.12.010

Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., Orlando, L., 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics 29, 1682–1684. https://doi.org/10.1093/bioinformatics/btt193

Jugovic, J., Črne, M., Lužnik, M., 2017. Movement, demography and behaviour of a highly mobile species: A case study of the black-veined white, Aporia crataegi (Lepidoptera: Pieridae). EJE 114, 113–122. https://doi.org/10.14411/eje.2017.016

Karam-Gemael, M., Decker, P., Stoev, P., Marques, M.I., Jr, A.C., 2020. Conservation of terrestrial invertebrates: a review of IUCN and regional Red Lists for Myriapoda. ZooKeys 930, 221–229. https://doi.org/10.3897/zookeys.930.48943

Kardos, M., Taylor, H.R., Ellegren, H., Luikart, G., Allendorf, F.W., 2016. Genomics advances the study of inbreeding depression in the wild. Evol Appl 9, 1205–1218. https://doi.org/10.1111/eva.12414

Keightley, P.D., Pinharanda, A., Ness, R.W., Simpson, F., Dasmahapatra, K.K., Mallet, J., Davey, J.W., Jiggins, C.D., 2015. Estimation of the Spontaneous Mutation Rate in Heliconius melpomene. Molecular Biology and Evolution 32, 239–243. https://doi.org/10.1093/molbev/msu302

Khan, A., Patel, K., Shukla, H., Viswanathan, A., van der Valk, T., Borthakur, U., Nigam, P., Zachariah, A., Jhala, Y.V., Kardos, M., Ramakrishnan, U., 2021. Genomic evidence for inbreeding depression and purging of deleterious genetic variation in Indian tigers. Proceedings of the National Academy of Sciences 118, e2023018118. https://doi.org/10.1073/pnas.2023018118

Kim, M.J., Cho, Y., Wang, A.R., Kim, S.-S., Choi, S.-W., Kim, I., 2020. Population genetic characterization of the black-veined white, Aporia crataegi (Lepidoptera: Pieridae), using novel microsatellite markers and mitochondrial DNA gene sequences. Conserv Genet 21, 359–371. https://doi.org/10.1007/s10592-020-01257-7

Kim, T., Han, Y.-G., Kwon, O., Cho, Y., 2015. Changes in Aporia crataegi's potential habitats in accordance with climate changes in the northeast Asia. Journal of Ecology and Environment 38, 15–23. https://doi.org/10.5141/ecoenv.2015.002

Korlević, P., McAlister, E., Mayho, M., Makunin, A., Flicek, P., Lawniczak, M.K.N., 2021. A Minimally Morphologically Destructive Approach for DNA Retrieval and Whole-Genome Shotgun Sequencing of Pinned Historic Dipteran Vector Species. Genome Biol Evol 13, evab226. https://doi.org/10.1093/gbe/evab226

Korneliussen, T.S., Albrechtsen, A., Nielsen, R., 2014. ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics 15, 356. https://doi.org/10.1186/s12859-014-0356-4

Kutschera, V.E., Kierczak, M., van der Valk, T., von Seth, J., Dussex, N., Lord, E., Dehasque, M., Stanton, D.W.G., Khoonsari, P.E., Nystedt, B., Dalén, L., Díez-del-Molino, D., 2022. GenErode: a bioinformatics pipeline to investigate genome erosion in endangered and extinct species. BMC Bioinformatics 23, 228. https://doi.org/10.1186/s12859-022-04757-0

Kyriazis, C.C., Beichman, A.C., Brzeski, K.E., Hoy, S.R., Peterson, R.O., Vucetich, J.A., Vucetich, L.M., Lohmueller, K.E., Wayne, R.K., 2023. Genomic Underpinnings of Population Persistence in Isle Royale Moose. Molecular Biology and Evolution 40, msad021. https://doi.org/10.1093/molbev/msad021

Lenth, R., 2023. emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.8.7.

Li, H., Durbin, R., 2011. Inference of human population history from individual whole-genome sequences. Nature 475, 493–496. https://doi.org/10.1038/nature10231

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Liu, N., Hargreaves, V.V., Zhu, Q., Kurland, J.V., Hong, J., Kim, W., Sher, F., Macias-Trevino, C., Rogers, J.M., Kurita, R., Nakamura, Y., Yuan, G.-C., Bauer, D.E., Xu, J., Bulyk, M.L., Orkin, S.H., 2018. Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch. Cell 173, 430-442.e17. https://doi.org/10.1016/j.cell.2018.03.016

Liu, S., Westbury, M.V., Dussex, N., Mitchell, K.J., Sinding, M.-H.S., Heintzman, P.D., Duchêne, D.A., Kapp, J.D., von Seth, J., Heiniger, H., Sánchez-Barreiro, F., Margaryan, A., André-Olsen, R., De Cahsan, B., Meng, G., Yang, C., Chen, L., van der Valk, T., Moodley, Y., Rookmaaker, K., Bruford, M.W., Ryder, O., Steiner, C., Bruins-van Sonsbeek, L.G.R., Vartanyan, S., Guo, C., Cooper, A., Kosintsev, P., Kirillova, I., Lister, A.M., Marques-Bonet, T., Gopalakrishnan, S., Dunn, R.R., Lorenzen, E.D., Shapiro, B., Zhang, G., Antoine, P.-O., Dalén, L., Gilbert, M.T.P., 2021. Ancient and modern genomes unravel the evolutionary history of the rhinoceros family. Cell 184, 4874-4885.e16. https://doi.org/10.1016/j.cell.2021.07.032

Lohse, K., 2023. The genome sequence of the Common Blue, Polyommatus icarus (Rottemburg, 1775). [version 1; peer review: 1 approved]. (https://doi.org/10.12688/wellcomeopenres.18772.1). Wellcome Open Res 8:72.

Lucena-Perez, M., Kleinman-Ruiz, D., Marmesat, E., Saveljev, A.P., Schmidt, K., Godoy, J.A., 2021. Bottleneck-associated changes in the genomic landscape of genetic diversity in wild lynx populations. Evol Appl 14, 2664–2679. https://doi.org/10.1111/eva.13302

Lynch, M., Conery, J., Bürger, R., 1995a. Mutational meltdowns in sexual populations. Evolution 49, 1067–1080. https://doi.org/10.1111/j.1558-5646.1995.tb04434.x

Lynch, M., Conery, J., Burger, R., 1995b. Mutation Accumulation and the Extinction of Small Populations. The American Naturalist 146, 489–518.

Mackintosh, A., Vila, R., Martin, S.H., Setter, D., Lohse, K., 2023. Do chromosome rearrangements fix by genetic drift or natural selection? Insights from Brenthis butterflies. Molecular Ecology n/a. https://doi.org/10.1111/mec.17146

Mackintosh, A., Laetsch, D. R., Hayward, A., Charlesworth, B., Waterfall, M., Vila, R., & Lohse, K. (2019). The determinants of genetic diversity in butterflies. *Nature Communications*, **10**, 3466. https://doi.org/10.1038/s41467-019-11308-4

MacLachlan, R., 1893. The Decadence of British Butterflies. The Entomologist's Monthly Magazine 29, 132–138.

Mattila, A.L.K., Duplouy, A., Kirjokangas, M., Lehtonen, R., Rastas, P., Hanski, I., 2012. High genetic load in an old isolated butterfly population. Proceedings of the National Academy of Sciences 109, E2496–E2505. https://doi.org/10.1073/pnas.1205789109

McCulloch, G.A., Waters, J.M., 2023. Rapid adaptation in a fast-changing world: Emerging insights from insect genomics. Glob Chang Biol 29, 943–954. https://doi.org/10.1111/gcb.16512

Meisner, J., Albrechtsen, A., 2018. Inferring Population Structure and Admixture Proportions in Low-Depth NGS Data. Genetics 210, 719–731. https://doi.org/10.1534/genetics.118.301336

Middleton-Welling, J., Dapporto, L., García-Barros, E., Wiemers, M., Nowicki, P., Plazio, E., Bonelli, S., Zaccagno, M., Šašić, M., Liparova, J., Schweiger, O., Harpke, A., Musche, M., Settele, J., Schmucki, R., Shreeve, T., 2020. A new comprehensive trait database of European and Maghreb butterflies, Papilionoidea. Sci Data 7, 351. https://doi.org/10.1038/s41597-020-00697-7

Mullin, V.E., Stephen, W., Arce, A.N., Nash, W., Raine, C., Notton, D.G., Whiffin, A., Blagderov, V., Gharbi, K., Hogan, J., Hunter, T., Irish, N., Jackson, S., Judd, S., Watkins, C., Haerty, W., Ollerton, J., Brace, S., Gill, R.J., Barnes, I., 2023. First large-scale quantification study of DNA

preservation in insects from natural history collections using genome-wide sequencing. Methods Ecol Evol 14, 360–371. https://doi.org/10.1111/2041-210X.13945

Nakamura, Y., 2011. Conservation of butterflies in Japan: status, actions and strategy. J Insect Conserv 15, 5–22. https://doi.org/10.1007/s10841-010-9299-x

Näsvall, K., Boman, J., Talla, V., Backström, N., 2023. Base Composition, Codon Usage, and Patterns of Gene Sequence Evolution in Butterflies. Genome Biology and Evolution 15, evad150. https://doi.org/10.1093/gbe/evad150

Neaves, L.E., Eales, J., Whitlock, R., Hollingsworth, P.M., Burke, T., Pullin, A.S., 2015. The fitness consequences of inbreeding in natural populations and their implications for species conservation – a systematic map. Environmental Evidence 4, 5. https://doi.org/10.1186/s13750-015-0031-x

Neukamm, J., Peltzer, A., Nieselt, K., 2021. DamageProfiler: fast damage pattern calculation for ancient DNA. Bioinformatics 37, 3652–3653. https://doi.org/10.1093/bioinformatics/btab190

O'Brien, D., Laikre, L., Hoban, S., Bruford, M. W., Ekblom, R., Fischer, M. C., et al. (2022). Bringing together approaches to reporting on within species genetic diversity. *J. Appl. Ecol.* doi:10.1111/1365-2664.14225

Palkopoulou, E., Mallick, S., Skoglund, P., Enk, J., Rohland, N., Li, H., Omrak, A., Vartanyan, S., Poinar, H., Götherström, A., Reich, D., Dalén, L., 2015. Complete genomes reveal signatures of demographic and genetic declines in the woolly mammoth. Curr Biol 25, 1395–1400. https://doi.org/10.1016/j.cub.2015.04.007

Pratt, C., 1983. A modern review of the demise of Aporia crataegi L.: the black veined white. The entomologist's record and journal of variation 95, 161--166.

Preston, S.D., Liao, J.D., Toombs, T.P., Romero-Canyas, R., Speiser, J., Seifert, C.M., 2021. A case study of a conservation flagship species: the monarch butterfly. Biodivers Conserv 30, 2057–2077. https://doi.org/10.1007/s10531-021-02183-x

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., Sham, P.C., 2007. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am J Hum Genet 81, 559–575.

R Core Team, 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rhie, A., McCarthy, S.A., Fedrigo, O., Damas, J., Formenti, G., Koren, S., Uliano-Silva, M., Chow, W., Fungtammasan, A., Kim, J., Lee, C., Ko, B.J., Chaisson, M., Gedman, G.L., Cantin, L.J., Thibaud-Nissen, F., Haggerty, L., Bista, I., Smith, M., Haase, B., Mountcastle, J., Winkler, S., Paez, S., Howard, J., Vernes, S.C., Lama, T.M., Grutzner, F., Warren, W.C., Balakrishnan, C.N., Burt, D., George, J.M., Biegler, M.T., Iorns, D., Digby, A., Eason, D., Robertson, B., Edwards, T., Wilkinson, M., Turner, G., Meyer, A., Kautt, A.F., Franchini, P., Detrich, H.W., Svardal, H., Wagner, M., Naylor, G.J.P., Pippel, M., Malinsky, M., Mooney, M., Simbirsky, M., Hannigan, B.T., Pesout, T., Houck, M., Misuraca, A., Kingan, S.B., Hall, R., Kronenberg, Z., Sović, I., Dunn, C., Ning, Z., Hastie, A., Lee, J., Selvaraj, S., Green, R.E., Putnam, N.H., Gut, I., Ghurye, J., Garrison, E., Sims, Y., Collins, J., Pelan, S., Torrance, J., Tracey, A., Wood, J., Dagnew, R.E., Guan, D., London, S.E., Clayton, D.F., Mello, C.V., Friedrich, S.R., Lovell, P.V., Osipova, E., Al-Ajli, F.O., Secomandi, S., Kim, H., Theofanopoulou, C., Hiller, M., Zhou, Y., Harris, R.S., Makova, K.D., Medvedev, P., Hoffman, J., Masterson, P., Clark, K., Martin, F., Howe, Kevin, Flicek, P., Walenz, B.P., Kwak, W., Clawson, H., Diekhans, M., Nassar, L., Paten, B., Kraus, R.H.S., Crawford, A.J., Gilbert, M.T.P., Zhang, G., Venkatesh, B., Murphy, R.W., Koepfli, K.-P., Shapiro, B., Johnson, W.E., Di Palma, F., Marques-Bonet, T., Teeling, E.C., Warnow, T., Graves, J.M., Ryder, O.A., Haussler, D., O'Brien, S.J., Korlach, J., Lewin, H.A., Howe, Kerstin, Myers, E.W., Durbin, R., Phillippy, A.M., Jarvis, E.D., 2021. Towards complete

and error-free genome assemblies of all vertebrate species. Nature 592, 737–746. https://doi.org/10.1038/s41586-021-03451-0

Robinson, J.A., Räikkönen, J., Vucetich, L.M., Vucetich, J.A., Peterson, R.O., Lohmueller, K.E., Wayne, R.K., 2019. Genomic signatures of extensive inbreeding in Isle Royale wolves, a population on the threshold of extinction. Sci. Adv. 5, eaau0757. https://doi.org/10.1126/sciadv.aau0757

Robinson, J., Kyriazis, C.C., Yuan, S.C., Lohmueller, K.E., 2023. Deleterious Variation in Natural Populations and Implications for Conservation Genetics. Annu Rev Anim Biosci 11, 93–114. https://doi.org/10.1146/annurev-animal-080522-093311

Rowe, G., Sweet, M., Beebee, T., 2017. An Introduction to Molecular Ecology, Third Edition. Oxford University Press, Oxford, New York.

Saccheri, I., Kuussaari, M., Kankare, M., Vikman, P., Fortelius, W., Hanski, I., 1998. Inbreeding and extinction in a butterfly metapopulation. Nature 392, 491–494. https://doi.org/10.1038/33136

Sarabia, C., vonHoldt, B., Larrasoaña, J.C., Uríos, V., Leonard, J.A., 2021. Pleistocene climate fluctuations drove demographic history of African golden wolves (Canis lupaster). Molecular Ecology 30, 6101–6120. https://doi.org/10.1111/mec.15784

Schmoldt, A., Benthe, H.F., Haberland, G., 1975. Digitoxin metabolism by rat liver microsomes. Biochem Pharmacol 24, 1639–1641.

Solonkin, I.A., Shkurikhin, A.O., Oslina, T.S., Zakharova, E.Y., 2021. Changes in the body size of black-veined white, Aporia crataegi (Lepidoptera: Pieridae), recorded in a natural population in response to different spring weather conditions and at different phases of an outbreak. Eur. J. Entomol.

South, R., 1906. The Butterflies of the British Isles. Frederick Warne.

Teixeira, J.C., Huber, C.D., 2021. The inflated significance of neutral genetic diversity in conservation genetics. Proc Natl Acad Sci U S A 118, e2015096118. https://doi.org/10.1073/pnas.2015096118

Thomas, J. A., 2005. Monitoring change in the abundance and distribution of insects using butterflies and other indicator groups. Philosophical Transactions of the Royal Society B: Biological Sciences 360, 339–357. https://doi.org/10.1098/rstb.2004.1585

Tscharntke, T., Steffan-Dewenter, I., Kruess, A., Thies, C., 2002. Characteristics of insect populations on habitat fragments: A mini review. Ecol Res 17, 229–239. https://doi.org/10.1046/j.1440-1703.2002.00482.x

van der Auwera, G., O'Connor, B.D., 2020. Genomics in the Cloud: Using Docker, GATK, and WDL in Terra. O'Reilly Media, Incorporated.

van der Valk, T., Díez-Del-Molino, D., Marques-Bonet, T., Guschanski, K., Dalén, L., 2019. Historical Genomes Reveal the Genomic Consequences of Recent Population Decline in Eastern Gorillas. Curr Biol 29, 165-170.e6. https://doi.org/10.1016/j.cub.2018.11.055

van Swaay, C., Van Strien, A., Aghababyan, K., Åström, S., Botham, M., Brereton, T., Carlisle, B., Chambers, P., Collins, S., Dopagne, C., Jiménez, R., Feldmann, R., Fernández-García, J., Fontaine, B., Goloshchapova, S., Gracianteparaluceta, A., Harpke, A., Heliölä, J., Khanamirian, G., Warren, M., 2016. The European Butterfly Indicator for Grassland species: 1990-2015. https://doi.org/10.13140/RG.2.2.31912.37127

Wagner, D.L., Grames, E.M., Forister, M.L., Berenbaum, M.R., Stopak, D., 2021. Insect decline in the Anthropocene: Death by a thousand cuts. Proceedings of the National Academy of Sciences 118, e2023989118. https://doi.org/10.1073/pnas.2023989118

Warren, M.S., 1992. The Conservation of British Butterflies, in: The Ecology of Butterflies in Britain, R.L.H. Dennis (Ed.). pp. 246–274.

Warren, M.S., Maes, D., van Swaay, C.A.M., Goffart, P., Van Dyck, H., Bourn, N.A.D., Wynhoff, I., Hoare, D., Ellis, S., 2021. The decline of butterflies in Europe: Problems, significance, and possible solutions. Proc Natl Acad Sci U S A 118, e2002551117. https://doi.org/10.1073/pnas.2002551117

Watson, R.T., Baste, I.A., Larigauderie, A., Leadley, P., Pascual, U., Baptiste, B., Demissew, S., Dziba, L., Erpul, G., Fazel, A.M., Fischer, M., Hernández, A.M., Karki, M., Mathur, V., Pataridze, T., Pinto, I.S., Stenseke, M., Török, K., Vilá, B., 2019. IPBES Global Assessment Report on Bioversity and Ecosystem Services.

Webster, M.T., Beaurepaire, A., Neumann, P., Stolle, E., 2023. Population Genomics for Insect Conservation. Annual Review of Animal Biosciences 11, 115–140. https://doi.org/10.1146/annurev-animal-122221-075025

Willoughby, J.R., Sundaram, M., Wijayawardena, B.K., Kimble, S.J.A., Ji, Y., Fernandez, N.B., Antonides, J.D., Lamb, M.C., Marra, N.J., DeWoody, J.A., 2015. The reduction of genetic diversity in threatened vertebrates and new recommendations regarding IUCN conservation rankings. Biological Conservation 191, 495–503. https://doi.org/10.1016/j.biocon.2015.07.025

Wood, D.E., Lu, J., Langmead, B., 2019. Improved metagenomic analysis with Kraken 2. Genome Biology 20, 257. https://doi.org/10.1186/s13059-019-1891-0

Xue, Y., Prado-Martinez, J., Sudmant, P.H., Narasimhan, V., Ayub, Q., Szpak, M., Frandsen, P., Chen, Y., Yngvadottir, B., Cooper, D.N., de Manuel, M., Hernandez-Rodriguez, J., Lobon, I., Siegismund, H.R., Pagani, L., Quail, M.A., Hvilsom, C., Mudakikwa, A., Eichler, E.E., Cranfield, M.R., Marques-Bonet, T., Tyler-Smith, C., Scally, A., 2015. Mountain gorilla genomes reveal

the impact of long-term population decline and inbreeding. Science 348, 242–245. https://doi.org/10.1126/science.aaa3952

Figure Legends

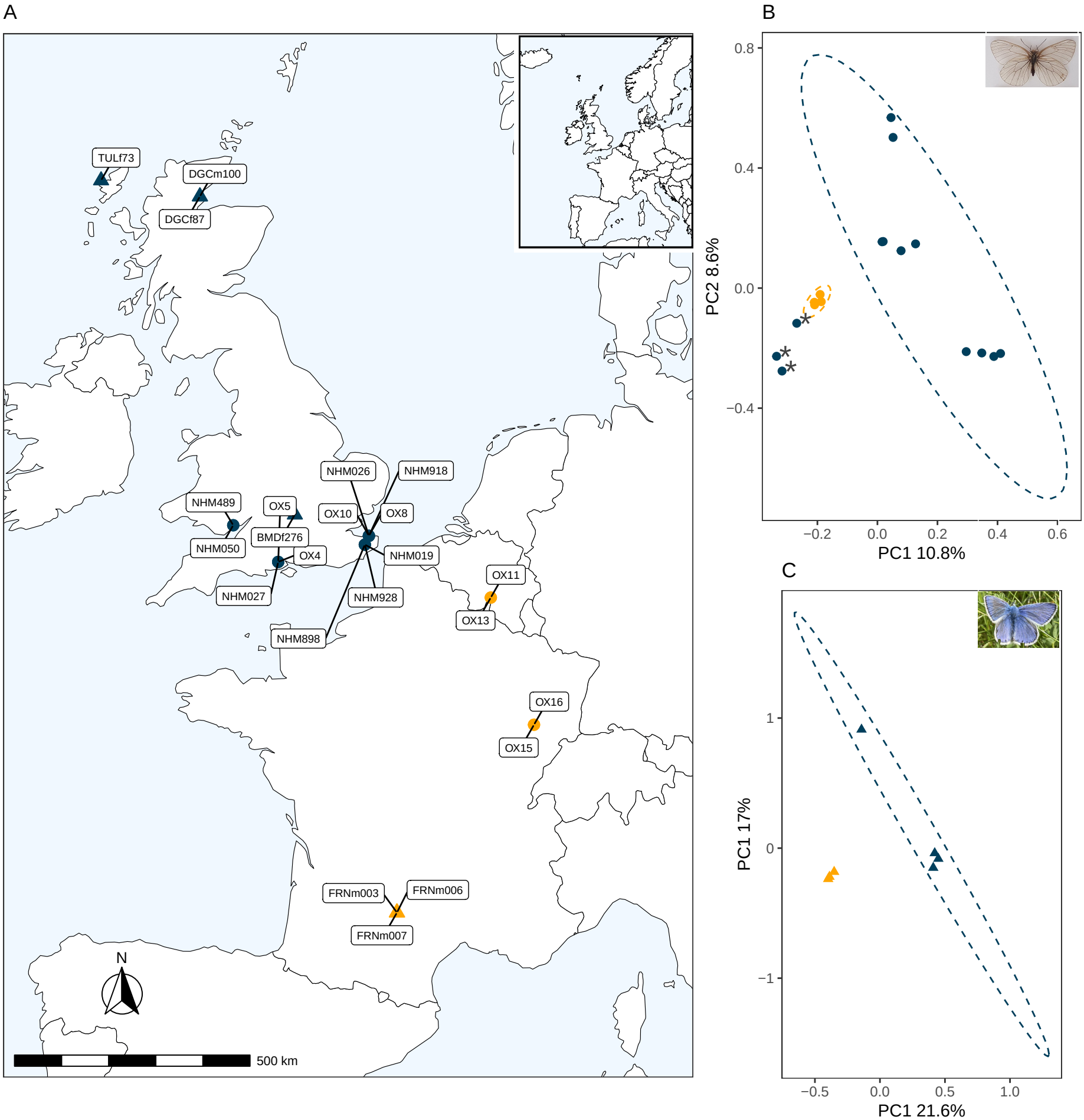
Figure 1 (A): Collection locations (based on museum records) of historical A. crataegi *specimens (circles) and of modern* P. icarus *(triangles). NHM prefix refers to specimens obtained from Natural History Museum, London and OX for those obtained from Oxford University Natural History Museum, Oxford. Inset: photograph of pinned A. crataegi museum specimen (B): Principal component analysis (PCA) for* A. crataegi *SNPs. The PCA is based on 14,930 LD-pruned SNPs based on hard genotype calls. A PCA using genotype likelihoods shows a similar pattern (Fig S3). Asterisks indicate UK-collected specimens that have clustered with European specimens. (C): PCA for* P. icarus *specimens based on 197,182 LD-pruned SNPs showing similar differentiation between British and European specimens. Inset images: A pinned* A. crataegi *specimen and live* P. icarus *in their respective plots.*

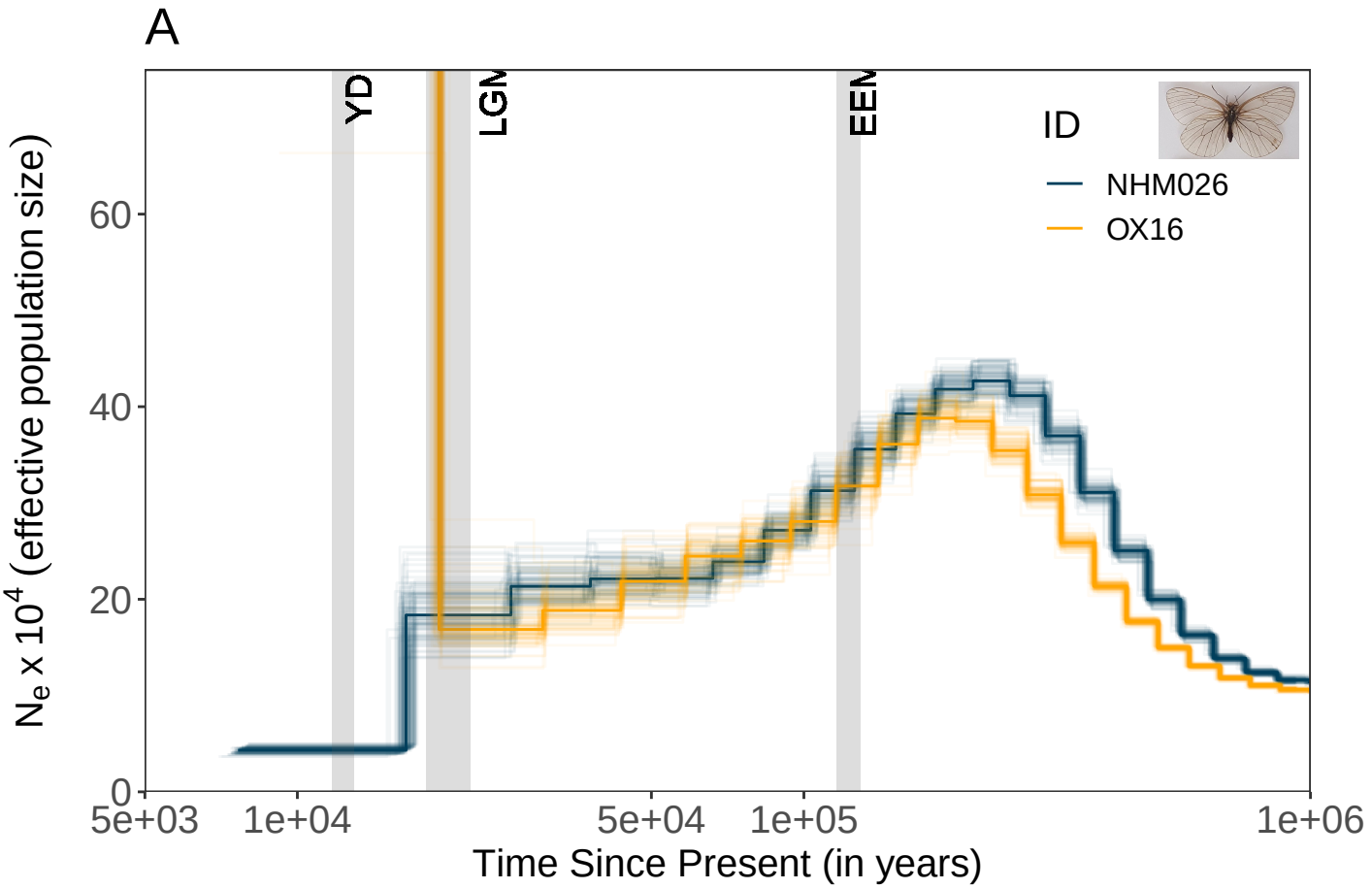
Figure 2: (A): PSMC plot for A. crataegi *specimens, showing changing* N_e *over time and population divergence between a GB specimen (blue curve) and a European specimen (yellow curve). All A. crataegi specimens PSMC curves are shown in S6. (B): PSMC for modern* P. icarus *specimens, with a GB specimen (blue curve) and a European specimen (yellow curve). Divergence timing is shown as before the LGM due to interval settings during the PSMC run (-p). YD = Younger Dryas, LGM= Last Glacial Maximum, EEM= Eemian Interglacial Period*

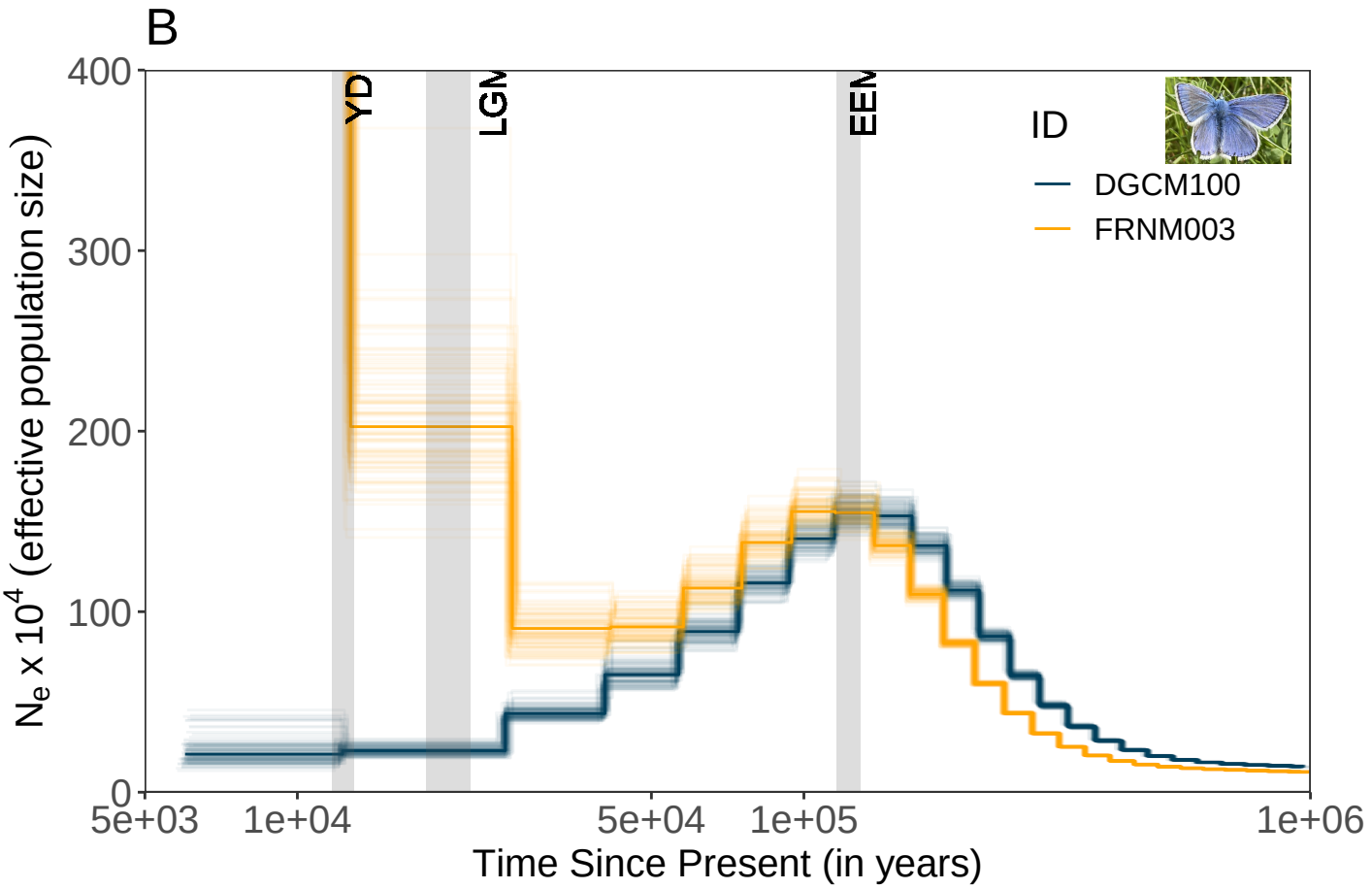
*Figure 3: (A): Individual heterozygosity rates in A. crataegi specimens from Europe and GB, calculated using ANGSD. Dashed lines show heterozygosity difference between Europe and GB. Significance codes: n.s. = not significant, * = p <0.05, ** = p <0.01, *** = p < 0.001. (B): Individual heterozygosity rates in* P.icarus *specimens from Europe and GB (C): RoH in* A. crataegi *specimens from Europe and GB. Open bars show the total proportion of the genome in RoH ≥ 100 KB, light grey bars show RoH ≥ 500kb and dark grey bars show proportions in RoH ≥ 1 MB. (D): RoH in* P.icarus *European and GB specimens. Dashed lines in C and D show significance of difference between European and GB specimens.*

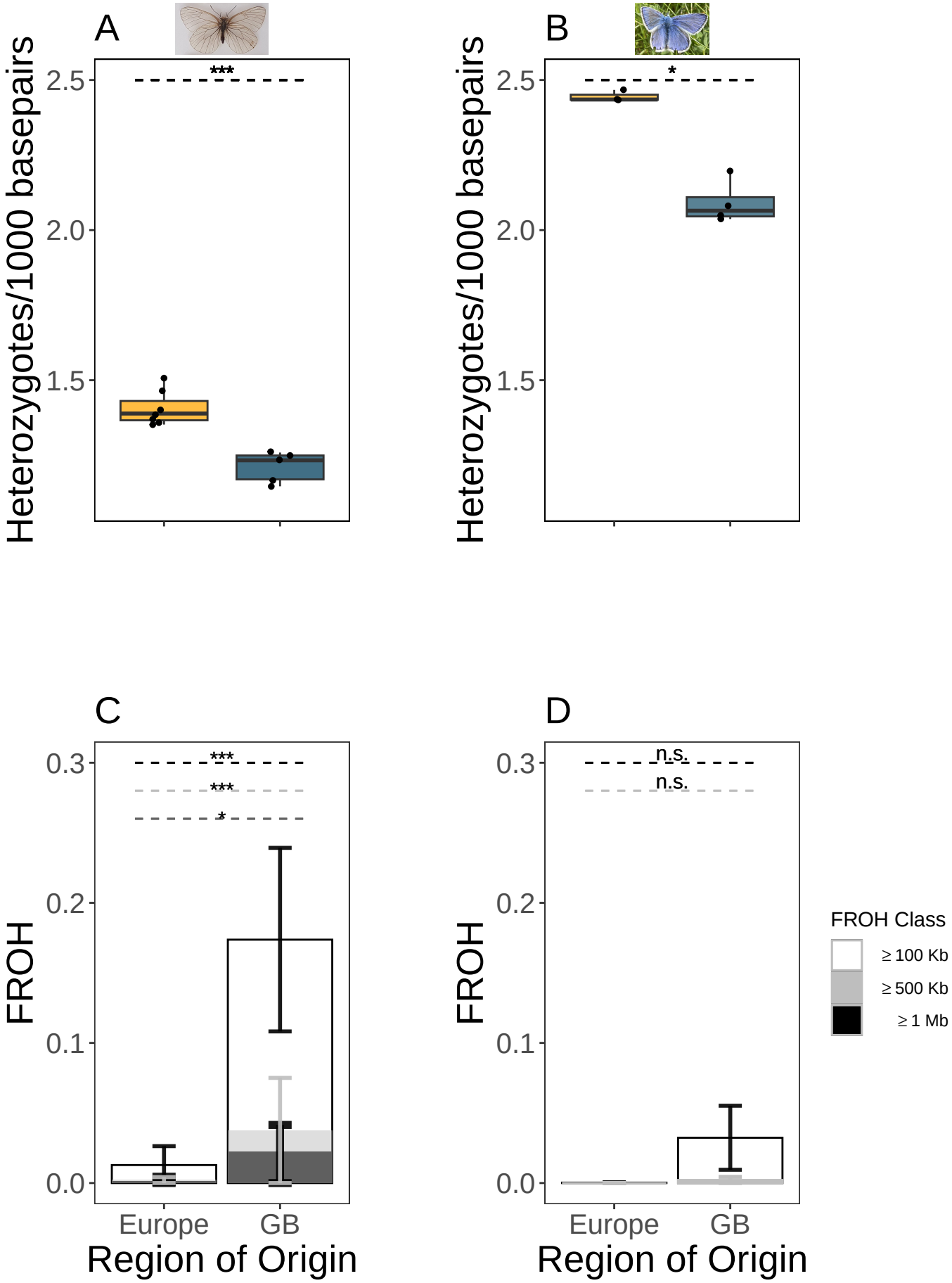
Figure. 4: (A-C): Total LoF, Missense and Synonymous variants found using snpEff in historic A. crataegi *(A.c.) specimens. (D-F): Total LoF, Missense and Synonymous variants found in* P. icarus*. Dashed lines show difference between Europe and GB. Significance codes: n.s. = not significant, * = p <0.05, ** = p <0.01, *** = p < 0.001. A.c: linear model with average coverage as a covariate*, *P.i. Welch's two sample t-test.*

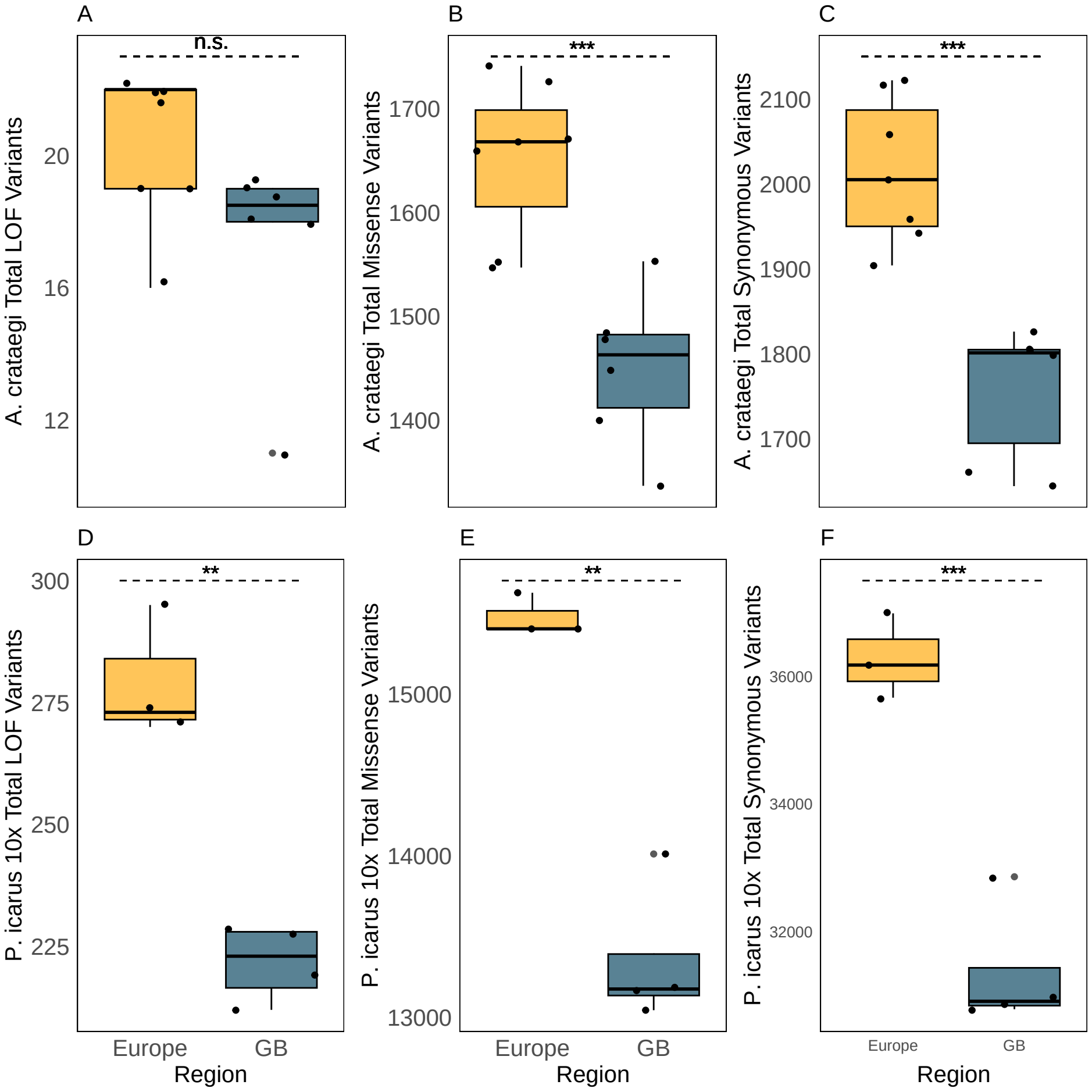
Figure 5 (A): R_{xy} ratio of shared derived alleles classified Synonymous, Missense and LoF variants in A. crataegi. *Red whiskers represent 95% confidence intervals based on jackknife estimates over chromosomes.A ratio of >1 represents an increase in frequency of alleles in population X (Gb) relative to population Y (Europe), 1 reflects no change and <1 a decrease (B) R_{xy} ratio of shared derived alleles of each class in* P. icarus.

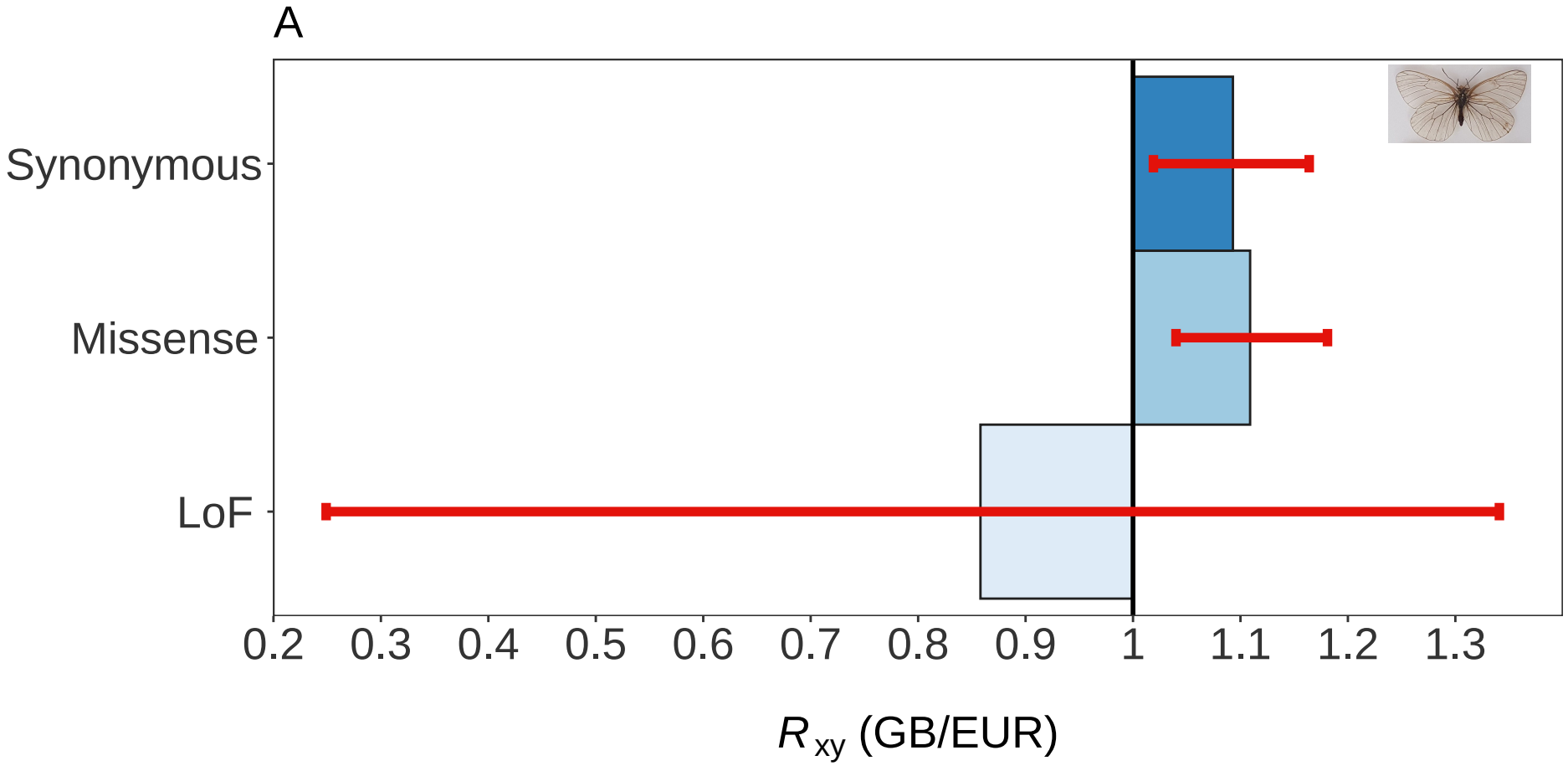

A
TULf73
DGCm100
DGCf87
NHM026
NHM918
NHM489
OX5
OX10
OX8
BMDf276
NHM050
OX4
NHM019
OX11
NHM027
NHM928
OX13
NHM898
OX16
OX15
FRNm003
FRNm006
FRNm007
N
500 km
B
PC2 8.6%
PC1 10.8%
0.8
0.4
0.0
−0.4
−0.2
0.0
0.2
0.4
0.6
C
PC1 17%
PC1 21.6%
1
0
−1
−0.5
0.0
0.5
1.0

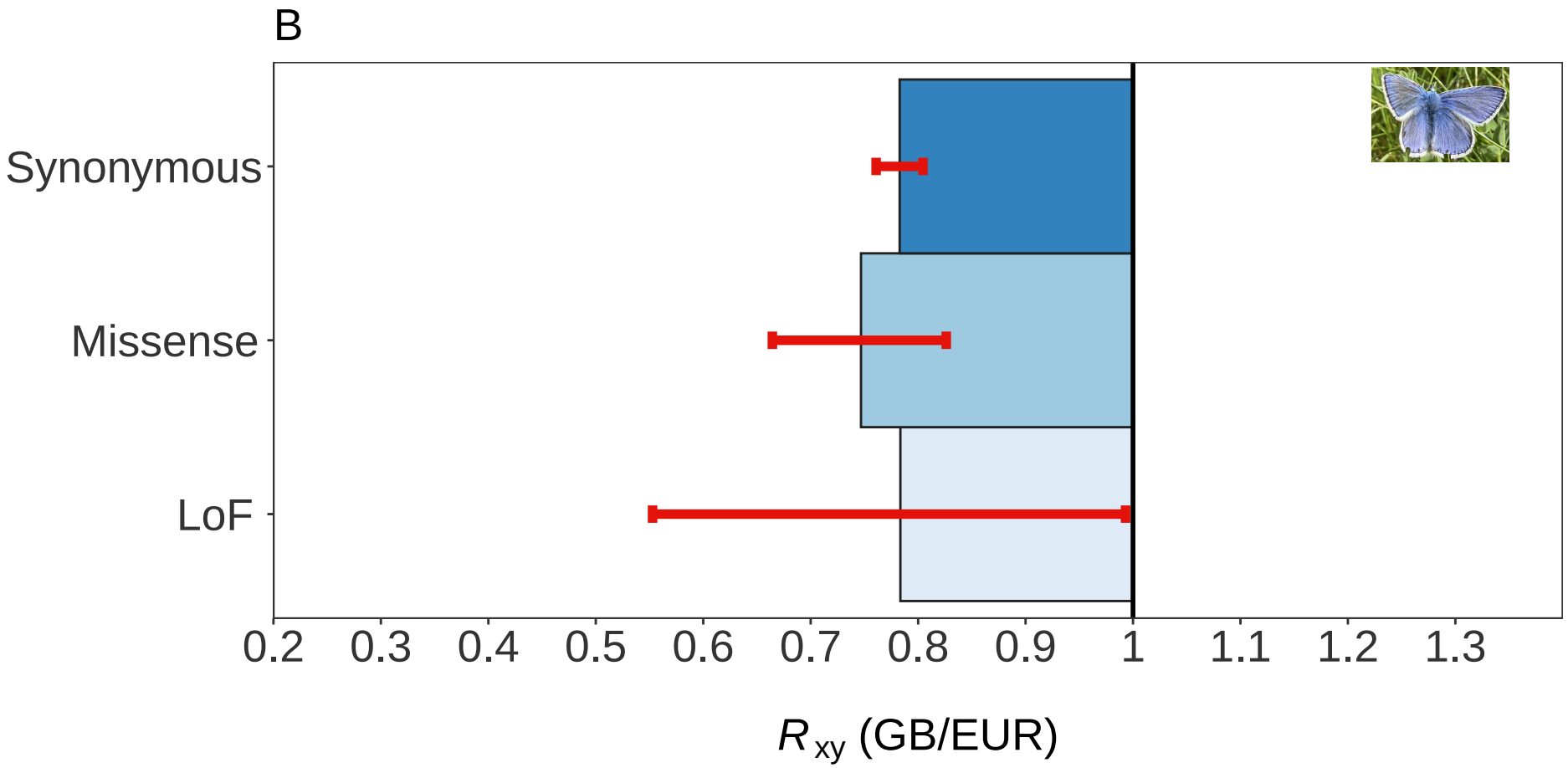

A
YD
LGM
EEM
ID
NHM026
OX16
N_e x 10^4 (effective population size)
0
20
40
60
5e+03
1e+04
5e+04
1e+05
1e+06
Time Since Present (in years)


B
YD
LGM
EEM
ID
DGCM100
FRNM003
N_e x 10^4 (effective population size)
0
100
200
300
400
5e+03
1e+04
5e+04
1e+05
1e+06
Time Since Present (in years)


A
B
C
D
Heterozygotes/1000 basepairs
2.5
2.0
1.5
***
*
FROH
0.3
0.2
0.1
0.0
n.s.
Europe
GB
Region of Origin
FROH Class
≥ 100 Kb
≥ 500 Kb
≥ 1 Mb


A
n.s.
A. crataegi Total LOF Variants
20
16
12
B
***
A. crataegi Total Missense Variants
1700
1600
1500
1400
C
***
A. crataegi Total Synonymous Variants
2100
2000
1900
1800
1700
D
**
P. icarus 10x Total LOF Variants
300
275
250
225
Europe
GB
Region
E
**
P. icarus 10x Total Missense Variants
15000
14000
13000
Europe
GB
Region
F
***
P. icarus 10x Total Synonymous Variants
36000
34000
32000
Europe
GB
Region


A
Synonymous
Missense
LoF
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1.1
1.2
1.3
R_{xy} (GB/EUR)


B
Synonymous
Missense
LoF
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1.1
1.2
1.3
R_{xy} (GB/EUR)