High levels of inbreeding with spatial and host-associated structure in lice of an endangered freshwater seal

Running title: Inbreeding and population structure in seal lice

Stephany Virrueta Herrera[1,2], Kevin P. Johnson[1], Andrew D. Sweet[3], Eeva Ylinen[4], Mervi Kunnasranta[4,5], and Tommi Nyman[6]

[1] Illinois Natural History Survey, Prairie Research Institute, University of Illinois, Champaign, IL, 61820 USA

[2] Program in Ecology, Evolution, and Conservation, University of Illinois, Urbana, IL, USA

[3] Department of Biological Sciences, Arkansas State University, Jonesboro, AR, 72401 USA

[4] Department of Environmental and Biological Sciences, University of Eastern Finland, P.O. Box 111, 80101 Joensuu, Finland

[5] Natural Resources Institute Finland, 80101 Joensuu, Finland

[6] Department of Ecosystems in the Barents Region, Svanhovd Research Station, Norwegian Institute of Bioeconomy Research, 9925 Svanvik, Norway

Corresponding author:

Tommi Nyman
Tommi.Nyman@nibio.no
Department of Ecosystems in the Barents Region
Svanhovd Research Station
Norwegian Institute of Bioeconomy Research
NO-9925 Svanvik
Norway

Abstract

Host-specialist parasites of endangered large vertebrates are in many cases more endangered than their hosts. In particular, low population densities and reduced among-host transmission rates are expected to lead to inbreeding within parasite infrapopulations living on single host individuals. Furthermore, spatial population structures of directly-transmitted parasites should be concordant with those of their hosts. Using population genomic approaches, we investigated inbreeding and population structure in a host-specialist seal louse (*Echinophthirius horridus*) infesting the Saimaa ringed seal (*Phoca hispida saimensis*), which is endemic to Lake Saimaa in Finland, and is one of the most endangered pinnipeds in the world. We conducted genome resequencing of pairs of lice collected from 18 individual Saimaa ringed seals throughout the Lake Saimaa complex. Our analyses showed high genetic similarity and inbreeding between lice inhabiting the same individual seal host, indicating low among-host transmission rates. Across the lake, genetic differentiation among individual lice was correlated with their geographic distance, and assignment analyses revealed a marked break in the genetic variation of the lice in the middle of the lake, indicating substantial population structure. These findings indicate that movements of Saimaa ringed seals across the main breeding areas of the fragmented Lake Saimaa complex may in fact be more restricted than suggested by previous population-genetic analyses of the seals themselves.

Keywords: conservation genomics, genome resequencing, host–parasite interactions, Saimaa ringed seal, seal louse

1 | INTRODUCTION

Many large vertebrates, particularly birds of prey and terrestrial and marine mammals, are listed as endangered because of habitat destruction, pollution, overexploitation, direct persecution, or climate change (Courchamp et al. 2018; IUCN 2021). Species belonging to the charismatic megafauna often act as flagship or umbrella species that attract public attention to the generally dire situation of natural ecosystems across the world (Berti et al. 2020; Thompson & Rog 2019). Endangered large vertebrates in many cases also constitute important model systems for studying the genetic effects of population bottlenecks and habitat fragmentation (Gousy-Leblanc et al. 2021; Luo et al. 2019; O'Brien et al. 2017). Conservation-genetic studies on endangered animals have focused on inbreeding and loss of genetic diversity (Karamanlidis et al. 2021; Rey-Iglesia et al. 2021), both of which can add to the direct threats imposed by reduced population size, such as Allee effects (Courchamp et al. 1999; Nagel et al. 2021) and sensitivity to environmental and demographic stochasticity (DeWoody et al. 2021; Díez-del-Molino et al. 2018; Kyriazis et al. 2021; Lande 1993; Spielman et al. 2004; Williams et al. 2021).

A fact often overlooked is that charismatic megafaunal species themselves constitute the habitat of other organisms. Large vertebrates host a multitude of ecto- and endoparasites, including lice, fleas, nematodes, and cestodes (Pérez et al. 2006; Thompson et al. 2018; Vlasman & Campbell 2004). In particular, highly host-specific parasites (*i.e.*, those found on only one host species) may be more endangered than their more obviously threatened hosts (Carlson et al. 2017; Dharmarajan et al. 2021; Dunn et al. 2009; Harris et al. 2014; Pérez et al. 2013; Rózsa & Vas 2015). While parasitic species are often small and visually unappealing to humans, they still constitute a substantial fraction of global biodiversity and an integral part of healthy ecosystems (Strona 2015; Thompson et al. 2018), and arguably have their own intrinsic value for ecosystem

function and nature conservation (Carlson et al. 2020; Gómez & Nichols 2013; Stork & Lyal 1993; Windsor 1997). Hence, preservation of parasite diversity is important for ensuring normal functioning of both ecosystem-level processes (Kwak et al. 2020; Milotic et al. 2020) and the immune defenses of their hosts (Spencer & Zuk 2016).

While conservation-genetic studies have predominantly focused on endangered large vertebrates, genetic investigations of their associated parasites are potentially highly useful for both fundamental and applied research. Two aspects are particularly important:

First, reduced population density of hosts will lower the availability of resources for parasites, and thereby diminish their chances for movements among host individuals. Transmission rates will be lowered especially for directly transmitted parasite species, *i.e.*, those that require close contact between host individuals for successful transmission. Reduced among-host transmission probabilities are expected to increase population-genetic structuring of parasites across host individuals (DiBlasi et al. 2018; Orsini et al. 2013; Sweet & Johnson 2018). Such effects should also be observed as elevated inbreeding within parasite infrapopulations inhabiting single host individuals (Detwiler & Criscione 2017).

Second, the spatial genetic structure of parasites can inform us about the population structures, movements, and social networks of their hosts (Gagne et al. 2022; Whiteman & Parker 2005). Directly transmitted host-specific parasites will in this respect again be most informative, because their genetic composition will in practice contain a record of past direct interactions among host individuals. From a research perspective, a practical benefit is that parasite genomes are often considerably smaller than those of their vertebrate hosts (de Moya et al. 2021; Kapusta et al. 2017; Zarowiecki & Berriman 2015). These smaller genomes make approaches that leverage genome sequencing for the collection of population-genomic data from many

individuals more cost effective (Johnson 2019). The short generation times and faster evolutionary rates in parasites may also mean that differences among subpopulations accumulate faster than in their hosts, potentially allowing analyses of population structuring across finer spatial and shorter temporal scales (Johnson et al. 2014; Martinů et al. 2020; Whiteman & Parker 2005).

Here, we investigated the levels of inbreeding and genetic differentiation in seal lice (*Echinophthirius horridus*) living on the endangered Saimaa ringed seal (*Pusa hispida saimensis*), with respect to both individual host and to geographic space (Fig. 1). The Saimaa ringed seal is a postglacial relic subspecies of the ringed seal and is endemic to Lake Saimaa in southern Finland (Fig. 2A). The current population of circa 400 individuals is in a slow recovery from a severe bottleneck in the 1980s, when seal numbers were down to less than 150 individuals (Kunnasranta et al. 2021). The postglacial isolation of nearly 10,000 years and the recent severe bottleneck have left their mark in the genetic composition of the Saimaa ringed seal population, which is one of the genetically most uniform pinniped populations on the Earth (Nyman et al. 2014; Palo et al. 2003; Peart et al. 2020; Stoffel et al. 2018). Microsatellite-based genetic analyses have shown that the fragmented shape of the Lake Saimaa complex (Fig. 2A), possibly in connection with the low population size, has led to population-genetic differentiation across the main breeding areas of the Saimaa ringed seal (Valtonen et al. 2012, 2014).

Seal lice are in many ways ideal for conservation-genetic analyses of endangered parasites on endangered hosts. They are obligate, strictly host-specific parasites that are directly transmitted among host individuals (Leidenberger et al. 2007). Louse genomes in general are small (100–200 Mbp) (Allen et al. 2017; Baldwin-Brown et al. 2021; de Moya et al. 2021), and their generation time is an order of magnitude shorter than those of seals (Kim 1975; Leonardi et al. 2013; Palo et

al. 2003). We estimated levels of genetic diversity and inbreeding, as well as the existence of host- and space-related genetic differentiation, in seal lice on Saimaa ringed seals by sampling pairs of lice from 18 seals across the entire Lake Saimaa complex (Fig. 2A). Based on phylogenomic and population-genomic datasets obtained through genome resequencing of the 36 sampled individuals, we investigated whether lice sampled from the same host individuals are on average more closely related than lice on different host individuals, and whether lice show signs of inbreeding on the population and host level. We also investigated whether genetic differences among lice are correlated with their geographic distances, and whether lice show differentiation across the main basins of the Lake Saimaa system. Finally, we contrasted the spatial genetic structures of the lice with results from prior population-genetic analyses of the Saimaa ringed seal (Valtonen et al. 2012, 2014, 2015).

2 | METHODS

2.1 | Sample collection

Pairs of lice were sampled from 18 seal individuals across Lake Saimaa through 2009–2017 (Fig. 2A, Supplementary Table S1). The sampling covers all major breeding areas of the Saimaa ringed seal, and the distance between the furthest samples is circa 150 km. Of the seals, nine were found dead (in the figures and tables below, these hosts are denoted by four-number codes), and nine were pups briefly captured for radio telemetry studies (below denoted by codes with two letters and two numbers) during long-term seal monitoring programs of the University of Eastern Finland and the Finnish Forest Management Authority (Metsähallitus). Telemetry studies have been approved by the local environmental authority Centre for Economic Development, Transport and the Environment (permit numbers: ESAELY/433/ 07.01/2012 and

ESA-2008-L-519-254) and the Animal Experiment Board in Finland (permit numbers: ESAVI/8269/04.10.07/2013 and ESAVI-2010-08380/Ym-23).

In addition to the 36 focal seal lice, two additional specimens were sampled and sequenced (Supplementary Table S1): individual Echor52, likewise from Lake Saimaa, was sequenced with higher coverage and was used for constructing target gene sequences for mapping reads of the focal lice (see below). The other non-focal specimen (Echor6) originated from a Ladoga ringed seal (*P. h. ladogensis*) and was used as an outgroup in phylogenomic analyses of the focal Lake Saimaa lice. All lice were collected into 99.5% ethanol in 2-ml screw-cap tubes and stored at –20°C, and each specimen was photographed as a voucher prior to DNA extraction.

2.2 | DNA extraction and genome sequencing

Whole lice were ground up individually in 1.5 mL Eppendorf tubes, and genomic DNA was isolated using the Qiagen QIAamp DNA Micro Kit (Qiagen, Valencia, CA, USA). The manufacturer's standard protocol was modified so that specimens were incubated in ATL buffer and proteinase K at 55 °C for 48 h instead of the recommended 1–3 h, as well as by substituting buffer AE with buffer EB (elution buffer). This was done to ensure maximal yield (greater than 5 ng) of DNA from the extraction. We quantified each DNA extract with a Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA).

Libraries for shotgun genomic sequencing were prepared from the extracts with Hyper Library construction kits (Kapa Biosystems, Wilmington, MA, USA). The libraries were quantified by qPCR and sequenced using 150 bp paired-end reads with an Illumina NovaSeq 6000 (Albany, New York). These libraries were multiplexed to consume approximately $1/96^{th}$ of a lane each, producing between 18.4 million and 145 million reads per library (Supplementary Table S1),

157 representing about 28X to 217X coverage assuming a 100 Mbp genome size. The Echor52
158 reference sample was multiplexed to consume 1/48th of a lane, producing over 148 million total
159 reads. The FASTQ files from the sequence data were generated and demultiplexed with bcl2fastq
160 v.2.20. All steps of library preparation, sequencing, and data file generation were carried out at
161 the Roy J. Carver Biotechnology Center (University of Illinois, Urbana, IL, USA). Raw reads
162 have been deposited in the NCBI GenBank SRA database (Supplementary Table S1).

163 **2.3 | Phylogenomic analyses**

164 To get an overview of relationships among the sampled seal lice, we constructed individual-level
165 phylogenomic trees based on sequences of 1107 single-copy protein-coding ortholog genes. For
166 this, we first used aTRAM (Allen et al. 2015, 2017) to assemble the protein-coding portions of
167 the focal ortholog genes based on amino acid target sequences from *Pediculus humanus*
168 (Johnson et al. 2013) and the 148 M total reads produced by the genomic sequencing library of
169 reference individual Echor52. We then mapped libraries of the focal lice to these assembled
170 target gene sequences from Echor52 using a reference-mapping pipeline script
171 (https://github.com/adsweet/louse_genomes/) and Bowtie2 (Langmead & Salzberg 2012). After
172 mapping, we sorted the BAM files and created pileup files using samtools v.1.7 (Li et al. 2009).
173 We used bcftools v.1.7 (Li et al. 2009) to call variants and to convert pileup files to VCF files.
174 Sites with sequence coverage less than 5X or greater than 100X, or with Phred quality scores
175 <28 were filtered using samtools. From these files, we created consensus sequences for each
176 gene from each individual louse using ambiguity coding for variants.

177 We aligned nucleotides across all individual lice for each gene separately using pasta v.1.8.2
178 (Mirarab et al. 2015). Using a custom Python script, we removed genes that contained fewer than
179 seven individuals, and then masked sites containing ≥40% gaps using trimal v.1.4 (Capella-

Gutiérrez et al. 2009). After filtering from the 1107-gene reference set, we were left with 1043 genes (with a total alignment length of 1,379,142 bp) that we used for phylogenomic analyses (Virrueta Herrera et al. 2022). With the aligned data, we performed both a phylogenetic analysis of the concatenated supermatrix and a coalescent analysis of gene trees to produce a species tree. All trees were rooted using the aforementioned louse specimen (Echor 6) collected from Ladoga ringed seal as an outgroup. For the concatenated method, we first ran our gene alignments through RAxML v.8.1.3 (Stamatakis 2014) and then used the resulting reduced alignment files to create a concatenated matrix. We performed a maximum likelihood (ML) analysis in RAxML, based on a GTR + Γ model of substitution and 100 rapid bootstrap replicates. For the coalescent analysis, we first estimated a tree for each gene alignment in RAxML using a GTR + Γ model, and then summarized the results of the gene-specific analyses as a coalescent species tree using ASTRAL v.4.10.6 (Mirarab et al. 2014), with quartet-based local posterior probability support for branches (Sayyari & Mirarab 2016).

2.4 | Population-genomic analyses

We constructed separate population-genomic datasets for estimation of genetic diversity, inbreeding, and population-genetic structuring due to seal host individuals and geographic location (Virrueta Herrera et al. 2022). First, we combined the individual VCF files from above into a single VCF file using the merge option in bcftools v. 1.7 (Li et al. 2009). We then ran the *populations* program in STACKS v. 2.5 (Rochette et al. 2019) to construct a Genepop-formatted file containing 2523 SNP sites for use in other population-genetic analysis programs.

We estimated standard population-level measures of genetic diversity (number of alleles, observed heterozygosity [H_O], heterozygosity within populations [H_S], total heterozygosity [H_T], and corrected heterozygosity [H'_T]) using the Genepop-formatted file in Genodive v.3.03

(Meirmans 2020). For the level of individual louse, we calculated the inbreeding coefficient (F) and standardized individual heterozygosity (Coltman et al. 1999) using the –het option in VCFtools v. 0.1.15 (Danecek et al. 2011) based on the combined VCF file. The number of sites that could be called as homozygous or heterozygous for individual lice ranged from 2984 to 3066 (mean = 3053.3; Supplementary Table S1).

To test whether the level of genetic diversity is correlated between lice from the same seals, we used mlRho v. 2.9 (Haubold et al. 2010) to calculate sample-specific mean theta (θ), which is defined as the population mutation rate, or $\theta = 4N_e\mu$, and which can be used as an indicator of heterozygosity and effective population size (Meyer et al. 2012). For this analysis, we converted pileup files generated from Bowtie2 to profile (.pro) files for each individual louse, and then ran mlRho with maximum distance (M) = 0. These files contained between 1,043,646 and 1,324,364 sites (mean = 1,279,978 sites; Supplementary Table S1). Finally, we plotted the mean θ of the two lice from each infrapopulation against each other and tested for any correlation between the estimates using reduced major axis regression in the *lmodel2* (Legendre 2018) package in R (R_Core_Team 2021). We also tested for an effect of lake area and seal host individual on mean θ using GLM ANOVA in IBM SPSS Statistics for Windows v. 27.0.1.0, with seal individual nested within lake area in the model.

We inferred the structuring effect of seal host individuals (*i.e.*, infrapopulation structure) by estimating genetic self-similarity and similarity among individual seal lice based on within- and between-individual kinship coefficients (Loiselle et al. 1995) in Genodive v.3.03 (Meirmans 2020). As a second estimate of differentiation among infrapopulations, we calculated overall F_{ST} in a dataset partitioned by seal host individual in Genodive.

We visualized overall genetic similarities among individual lice by Principal coordinates analysis (PCoA) in *adegenet* (Jombart 2008; Jombart & Ahmed 2011) in R. The PCoA method seeks the best approximation in reduced space of a matrix of Euclidean distances. Its principal components optimize the representation of the squared pairwise distances between individuals (Jombart 2016). We then assessed population structure by estimating the ancestry of individual lice using ADMIXTURE v.1.3 (Alexander et al. 2009). We ran ADMIXTURE for K (number of ancestral populations) = 1–10 with the cross-validation method to test for the optimal value of K. More optimal values of K will show lower cross-validation error relative to less optimal values.

To investigate spatial genetic differentiation in the seal louse population within Lake Saimaa, we used two methods:

First, we correlated genetic distances among louse individuals to their geographic distances. Because lice from the same seal cannot be considered independent replicates in an isolation-by-distance (IBD) analysis, we added the sampling-site coordinates of each individual into the Genepop file, but then split the file into ten separate datasets containing only one randomly selected louse per seal individual. The existence of IBD was then tested for each dataset in GenePop v.4.7.5 (Rousset 2008), based on genetic distances estimated based on the $\hat{a}$ statistic (Rousset 2000) and ln geographic distances estimated based on the sampling-site coordinates. Statistical significance of the regression slopes was inferred on the basis of 95% confidence intervals obtained through ABC bootstrapping (Leblois et al. 2003) and Mantel tests based on 10,000 permutations of individual locations. We also performed a corresponding analysis including all 36 lice, but with the minimum geographic distance among individuals set to 10^{-5}, so that lice from the same seal were not included in the estimation of the regression coefficient.

The second test for spatial effects was done with a hierarchical locus-by-locus AMOVA performed in Arlequin v. 3.5.2.2. (Excoffier & Lischer 2010). Prior to the analysis, we converted the Genepop-formatted data file to Arlequin format using the Widgetcon 1.0.0. website (Aydın et al. 2019) and manual editing. In the analysis, we divided the lice into three main areas (Northern Saimaa + Haukivesi, Pihlajavesi, and Southern Saimaa) defined based on the main basins and breeding areas of ringed seals within Lake Saimaa (Fig. 2A). This division scheme is slightly simplified from the one used in the analyses of spatial genetic differentiation in Saimaa ringed seals by Valtonen et al. (2012, 2014), because lice from only a single seal from Northern Saimaa were obtained, so we collapsed this sample of two lice into those from the adjacent Haukivesi population. The AMOVA was then performed with infrapopulations (seal host individuals) nested within lake area and including the level of louse individual in the analysis. Statistical significance of the effect of lake area and infrapopulation was determined by 10,000 permutations of seals (infrapopulations) among lake areas and lice among seals within areas.

3 | RESULTS

3.1 | Phylogenomic trees

The ML phylogeny based on the concatenated alignment revealed a few clear cases in which lice from the same seal host individual were each other's closest relatives (Supplementary Fig. S1A). The tree also showed some indication of lice from the same area being clustered close to each other, but bootstrap support values for groupings were generally very low across the tree, although this is not unexpected given these are individuals of the same species. Interestingly, by contrast, the coalescent ASTRAL tree revealed a clear structuring effect of seal host individual, with 14 out of 18 sampled louse infrapopulations coming out as monophyletic (*i.e.*, the two

individual lice from the same seal host individual were each other's closest relatives; Supplementary Fig. S1B). The general pattern of structuring by lake area was likewise more evident in the coalescent tree, although the support for the backbone structure of the phylogeny was weaker than for the clades formed by infrapopulations, which were for the most part strongly supported (Supplementary Fig. S1B).

3.2 | Population-genomic analyses

Overall observed heterozygosity in the focal seal louse population was 0.199 (s.d. 0.004), expected (total) heterozygosity 0.234 (s.d. 0.004), and corrected expected heterozygosity 0.238 (s.d. 0.004). Heterozygosity within infrapopulations was 0.164 (s.d. 0.003). On the level of individual lice, standardized heterozygosity ranged from 0.173 to 0.284 (Supplementary Table S1).

As expected, higher individual heterozygosity estimates corresponded to lower inbreeding coefficients (0.290 to -0.167) (Supplementary Table S1). Estimates of individual mean θ ranged between $4.68*10^{-4}$ and $7.99*10^{-4}$ (Supplementary Table S1) and were statistically significantly positively correlated between lice from the same infrapopulation (Fig. 3; reduced major axis regression $r = 0.556$; $P = 0.016$). A statistically significant effect of infrapopulation (nested within region) on mean θ was also revealed in the GLM ANOVA (df = 15, $F = 2.929$, $P = 0.016$), but estimates did not differ across the three regions of the lake (df = 2, $F = 2.046$, $P = 0.164$).

Between-individual kinship coefficients ranged between -0.205 and 0.478 and were generally highest between lice from the same infrapopulation (Supplementary Table S2). Self-similarities ranged between 0.463 and 0.709, with a mean of 0.585 (s.d. 0.062).

The overall F_{ST} among infrapopulations was 0.312, which was statistically highly significantly different from 0 ($P < 0.001$). The population structuring arising from the seal host individual is seen also in the PCoA ordination plot, in which lice from the same infrapopulation tended to cluster together (Fig. 2B).

The ADMIXTURE cross-validation analysis returned an optimal K value of 2 with a CV error value of 0.0012 (Supplementary Fig. S2). The analysis at the optimal $K = 2$ revealed a sharp change in ancestry proportions roughly in the middle of the lake, corresponding to the limit between Northern Saimaa + Haukivesi and the two southern parts of the lake (Fig. 2D).

Plotting the genetic distances among lice against their ln geographic distances revealed a classic IBD pattern (Fig. 2C). In the statistical analyses of the relationship using ten subsampled datasets, the mean intercept was 0.274 (range 0.178–0.323) and the mean slope parameter 0.045 (range 0.019–0.053). The bootstrapped 95% CI of the slope parameter did not include zero in any of the subsampled datasets (overall upper / lower range 0.012–0.063). P values estimated by Mantel tests were highly significant or significant at $P = 0.0002$–0.009 in nine cases and marginally significant at $P = 0.061$ in one case. For the analysis based on the complete dataset of 36 lice, the relationship between genetic and ln geographic distance was estimated as $\hat{a} = 0.224 + 0.043x$, with the 95% CI of the slope parameter being 0.035–0.050 and $P < 0.0001$ in the Mantel test.

The aforementioned patterns were largely summarized by the results of the hierarchical locus-by-locus AMOVA, which revealed statistically significant differentiation among the three lake areas as well as among infrapopulations within the areas (Table 1). The differentiation among lice within infrapopulations was strongly and statistically significantly negative, which is a further indication of inbreeding within populations of lice from the same seal individual (Table 1).

4 | DISCUSSION

Population-genetic investigations of endangered parasites can inform us about their population size, genetic diversity, and level of inbreeding, all of which have the potential to influence the likelihood of extinction through deterministic or stochastic processes (DeWoody et al. 2021; Kyriazis et al. 2021; Spielman et al. 2004). Importantly, detailed genetic surveys focusing on parasites of endangered vertebrates have wider applied value, as parasite-specific analyses can illuminate also biological features of their hosts and, thereby, aid in designing actions for conserving both the parasites *and* their hosts (Gagne et al. 2022; Sweet et al. 2020; Whiteman & Parker 2005). Here, we used genome-level data of seal lice living on the landlocked Saimaa ringed seal to gain insights into the population structure of the lice as well as their lake-endemic hosts. With a population of barely over 400 individuals, the Saimaa ringed seal is one of the most endangered pinniped populations on the Earth (Kunnasranta et al. 2021). Previous studies have shown that genetic diversity of the Saimaa ringed seal population is extremely low in comparison to other seal species (Martinez-Bakker et al. 2013; Nyman et al. 2014; Peart et al. 2020; Stoffel et al. 2018). In addition, the main breeding areas of the Lake Saimaa complex harbor partially isolated subpopulations (Valtonen et al. 2012, 2014, 2015). Given that the population density of seals within the lake is low and that seal lice require close contact between host individuals for transmission, we expected that lice inhabiting the same seal would tend to be closely related as a result of within-host inbreeding. Furthermore, we predicted that the low diversity and distinct spatial genetic structure found in the Saimaa ringed seals would be reflected in the genetic composition of their lice. Our phylogenomic and population-genomic analyses based on whole-genome resequencing data from 36 lice sampled from 18 seals across Lake Saimaa indeed supported all of these predictions. Below, we discuss our main results and their implications for

the conservation of Saimaa ringed seals, their host-specific lice, and endangered host–parasite systems in general.

4.1 | Genetic diversity, differentiation among infrapopulations, and inbreeding

The Saimaa ringed seal is genetically highly uniform in comparison to its sister subspecies in the Baltic Sea (*P. h. botnica*), Lake Ladoga, and the Arctic Ocean (*P. h. hispida*) (Martinez-Bakker et al. 2013; Nyman et al. 2014; Palo et al. 2003). The loss of diversity is apparently a consequence of a small founding population and long postglacial isolation (Nyman et al. 2014), as well as the severe anthropogenic 20th-century bottleneck (Peart et al. 2020; Stoffel et al. 2018). In our focal seal lice, overall heterozygosity estimates (H_O = 0.199, H'_T = 0.238) were not extremely low. In a study by DiBlasi et al. (2018), mean heterozygosities were 0.449 for pigeon body lice and 0.557 for wing lice. However, these latter values were based on microsatellite markers, which tend to have many alleles per locus and, hence, result in high heterozygosity estimates (Sunde et al. 2020). Our individual-level estimates of genomic diversity are, however, directly comparable to those of Leonardi et al. (2019), who used the same genomic markers to estimate θ values for five species of seal lice infesting Antarctic and Australian seals having very large population sizes. In their study, species-specific θ estimates based on individual lice ranged between 0.00107 and 0.00367, which is substantially higher than our individual-level estimates for *E. horridus* lice within Lake Saimaa (mean = 0.00062). The highest θ estimate in our dataset (0.00080) is also lower than the lowest values (range = 0.00087–0.00863) found by Sweet & Johnson (2018) for seven species of chewing lice on New World ground-doves. The level of genetic diversity of seal lice within Lake Saimaa therefore seems to directly reflect the low population size and genetic uniformity of their endangered hosts.

From the perspective of parasites of large vertebrates, each host individual constitutes a distinct resource "island" (Itescu 2019; Koop et al. 2014). If the frequency of among-host dispersal is low in relation to the generation time of the parasites, parasite populations on different host individuals (infrapopulations) will over time tend to become genetically differentiated from each other (DiBlasi et al. 2018; Huyse et al. 2005). Indeed, our phylogenomic trees, between-individual kinship coefficients, and estimates of among-infrapopulation F_{ST}'s consistently showed that lice collected from the same Saimaa ringed seal individual are on average genetically more similar than are individuals collected at random from the host population. Notably, the population-genetic differentiation found across lice collected from different seal individuals is not the only aspect that is affected by the fact that seal lice are distributed into distinct infrapopulations: Using the same specimens that were analyzed in this study, Doña et al. (2021) found that infrapopulation identity explained a major proportion of the variation in microbiome composition within individual lice.

The structuring imposed by infrapopulations is clearly visualized in the PCoA ordination, in which lice originating from the same seal are generally located close to each other (Fig. 2B). According to the hierarchical AMOVA controlling for within-lake spatial structure, variation among infrapopulations accounts for 21% of the genomic variation in the louse population (Table 1). Our estimated overall F_{ST} among infrapopulations (0.312) is high in comparison to studies on among-host differentiation in human body and head lice (F_{ST} = 0.048 in both; Leo et al. 2005), pigeon body (F_{ST} = 0.225) and wing (F_{ST} = 0.075) lice (DiBlasi et al. 2018), and feather lice on Galapagos hawks (pairwise F_{ST} = 0.145–0.183; Koop et al. 2014). Unfortunately, we can only make general comparisons among these different louse–host systems because the spatial scale of different studies varies considerably, and the high heterozygosity of

382 microsatellite markers used in previous studies will in theory suppress estimates of among-
383 population differentiation (Alcala & Rosenberg 2019; Jakobsson et al. 2013; Meirmans &
384 Hedrick 2011). Comparative studies have, however, indicated that microsatellite and SNP
385 markers produce roughly similar estimates of population differentiation (Lemopoulos et al. 2019;
386 Sunde et al. 2020). Hopefully, genomic approaches and the gene ortholog SNP based markers
387 applied here will in the future allow more direct comparisons of genetic variation and
388 differentiation measured from different study systems.

389 Based on previous studies of louse infrapopulations and the biology of seal louse transmission, it
390 was not unexpected to find genetic differentiation among infrapopulations of the Saimaa seal
391 lice. Despite their long coevolutionary history with aquatic mammals, seal lice are still
392 essentially terrestrial organisms (Leidenberger et al. 2007; Leonardi et al. 2013). Therefore,
393 transmission of lice requires direct contact between seals while they are not submerged in water
394 (Kim 1975). Within Lake Saimaa, lice are probably transmitted mainly between mothers and
395 pups during nursing (Fig. 1C) as is the case in other species of seal lice (*cf*. Kim 1975; Leonardi
396 et al. 2013). However, close seal-to-seal encounters occur also during the early-summer moulting
397 period, when two or more seals can share resting sites on large lakeside rocks (Biard et al. 2022).
398 Recent observations also indicate that multiple seals can co-inhabit the same resting lairs that the
399 seals dig into lakeside snowdrifts during winter (M. Kunnasranta, pers. obs.), so these may
400 provide additional opportunities for louse transmission.

401 In addition to the inbreeding caused by transmission dynamics, louse infrapopulations on single
402 seals are also presumably quite small, further increasing the level of inbreeding. In a sample of
403 49 seals in the collections of the University of Eastern Finland, the number of collected lice
404 ranged from one to 32. Seal lice are difficult to collect exhaustively and immature individuals

405 may go unnoticed, but it seems reasonable to assume that infrapopulation sizes range in the tens
406 rather than in the hundreds. Our sample of seal lice indeed showed clear genomic signs of
407 inbreeding. Individual *F* values are on average slightly positive and the hierarchical AMOVA
408 showed slightly negative estimates for differentiation between individuals from the same seal.
409 Furthermore, the mean of pairwise Loiselle's kinship coefficients within infrapopulations (0.31)
410 exceeds the expected value between parents and offspring or between siblings (0.25), and the
411 mean of self-similarity (0.58) likewise exceeds the expectation (0.50) in a randomly-mating
412 population. Interestingly, the level of genomic diversity and inbreeding varies among
413 infrapopulations, because estimates of θ, which is proportional to the effective population size
414 (Haubold et al. 2010), was found to be positively correlated between lice collected from the same
415 seal individual (Fig. 3). This variation evidently reflects substantial differences in
416 infrapopulation size and age, but potentially also stochastic immigration of unrelated individuals
417 into small and generally closed louse infrapopulations.

418 **4.2 | Spatial differentiation**

419 Spatial population-genetic differentiation in host-specific parasites is expected to be influenced
420 by the dispersal patterns of their host species, but spatial structuring can be either weaker or
421 stronger than in the hosts (Cole & Viney 2019; Dharmarajan et al. 2016; Mazé-Guilmo et al.
422 2016; McCoy et al. 2005; Sweet et al. 2020). Weaker differentiation is expected if the parasite
423 species also utilizes intermediate hosts or other host species, has a large effective population size
424 in relation to its host, or if it has a complex life cycle with a highly dispersive life stage (Blasco-
425 Costa & Poulin 2013; DiBlasi et al. 2018; Solórzano-García et al. 2021). By contrast, relatively
426 stronger differentiation is the norm if the parasite is host-specific, directly transmitted, occurs at

low prevalences, and has a comparatively short generation time and high mutation rate (Mazé-Guilmo et al. 2016).

Despite its large size, Lake Saimaa is in fact a labyrinthine watercourse system formed by several main basins connected by narrow straits (Fig. 2A). The fragmented structure of the lake has left its imprint in the genetic composition of the Saimaa ringed seal population, which exhibits an isolation-by-distance pattern and differences in the frequencies of mitochondrial haplotypes and nuclear microsatellite alleles across the main breeding areas (Valtonen et al. 2012, 2014, 2015). Similar to the patterns found in the seal hosts, our genome resequencing data of lice revealed a parallel isolation-by-distance gradient (Fig. 2C) and spatial differentiation (Fig. 2B) within Lake Saimaa. As a result, lice from all three main areas of our analysis tended to be grouped together in the PCoA ordination (Fig. 2B). Inspection of eigenvalues of the ordination axes additionally shows that most of the variation is explained by Axis 1, which largely corresponds to sampling locations in the north–south direction across the lake.

Importantly, our Admixture results reveal that the main division within the focal seal louse population occurs in the middle of the lake, around the Kyrönsalmi strait (Fig. 2A, D). Both shores of the strait are currently covered by the town of Savonlinna, with over 30,000 inhabitants. However, the area has had substantial human population at least since the foundation of the medieval St. Olaf's Castle on an island in the middle of the strait in 1475 (Taavitsainen 2005). Given that Saimaa ringed seals were actively hunted until their protection in 1955, the growing human population may have essentially stalled seal—and seal louse—migration between the northern and southern halves of the lake for some five to six hundreds of years. The low signature of northern genomic ancestry in five lice from Southern Saimaa (Fig. 2D) might conceivably result from the experimental translocation of a female seal (Phs152) from Haukivesi

to the southern parts of the lake in 1992. This move may have led to inadvertent north–south translocation of lice (and, hence, northern genetic variation), as seal Phs152 is known to have reproduced in its new home range, and it was still alive in 2020 (Kunnasranta et al. 2021).

It is noteworthy that the differentiation in seal lice (Fig. 2D) appears to be stronger than that estimated for their seal hosts on the basis of mtDNA and microsatellite data by Valtonen et al. (2012, 2014, 2015). The seal population exhibits statistically significant lake-wide differences in the frequencies mtDNA haplotypes and microsatellite alleles, but microsatellite-based assignment analyses by Valtonen et al. (2014) produced spatially restricted clusters only if sampling-site coordinates were used as background data (priors) in the analyses. In addition, the clusters were not strictly area-specific, so that individuals belonging to most clusters could be found in several areas of the lake. The stronger spatial signal in lice is most likely due to our much larger genome-level dataset, but also to the fact that seal lice can produce several generations per year (Kim 1975; Leonardi et al. 2013), while the generation time of ringed seals has been estimated at circa 11 years (Palo et al. 2001). Hence, the seal louse population will accumulate spatial genetic differences substantially faster than their seal hosts.

5 | CONCLUSIONS

Our phylogenomic and population-genomic analyses of host-specific ectoparasitic *E. horridus* seal lice from the lake-endemic and endangered Saimaa ringed seals show that the louse population consists of genetically distinct infrapopulations that differ among seal individuals and experience high levels of inbreeding. Furthermore, comparisons to genome-level studies from other louse groups suggest that overall genetic diversity within the focal seal louse population is low—a result that seems to parallel the genetic uniformity of the Saimaa ringed seal population

(Nyman et al. 2014; Palo et al. 2003). However, further studies are required for inferring the ecological and evolutionary relevance of reduced genetic diversity in the focal seal lice. While inbreeding and low genetic variation can suppress viability and reproductive success at the level of both individuals (Blomqvist et al. 2010; Kardos et al. 2016) and populations (Ekroth et al. 2019; Spielman et al. 2004), many parasites are known to experience regular cycles inbreeding due to their biological characteristics (Appelgren et al. 2018; Detwiler & Criscione 2017; Van Den Broeck et al. 2014). Hence, parasites may be tolerant to the negative effects of inbreeding (Price 1980), possibly through purging of deleterious genetic variation (Benesh et al. 2014). Inbred hosts have been shown to be more susceptible to parasitism in many species (Cassinello et al. 2001; Coltman et al. 1999; Hoffman et al. 2014), but far less is known about the effects of inbreeding on parasite performance (Forsman 2014; see also Benesh et al. 2014; Fredericksen et al. 2021). The endemic Saimaa ringed seals and their specialist lice therefore constitute a promising model system for investigating host susceptibility and parasite infectivity in a 'coevolutionary cold spot' in which interactions are highly specialized but in which both hosts and parasites have reduced genetic diversity.

Our population-genomic analyses revealed a distinct genetic discontinuity in the louse population at the Kyrönsalmi strait, which separates the northern and southern halves of the Lake Saimaa complex. Importantly, this division in the seal louse population suggests that the Saimaa ringed seals of the northern and southern parts of Lake Saimaa are more isolated from each other than mtDNA- and microsatellite-based analyses of the seals themselves have indicated. According to our data, the genetic effects may simply not yet have manifest in the seals due to their longer generation time. To make the comparisons between seals and their lice more comparable, the investigations based on mtDNA and microsatellites by Valtonen et al. (2012,

2014, 2015) should be followed up by genome-level analyses of the seal population in order to obtain a clear view of their spatial differentiation within the Lake Saimaa complex. Overall, our results highlight how genome-level analyses of parasites can provide a tractable, cost-effective, and sensitive early-warning system for detecting host population fragmentation before the genetic effects are evident in their vertebrate hosts.

Acknowledgements

We would like to thank especially the numerous seal researchers and conservation biologists who provided samples of hard-to-find seal lice: Marja Isomursu (Finnish Food Authority), Miina Auttila (Finnish Forest Management Authority (Metsähallitus)), and Marja Niemi, Vincent Biard, and other members of the Saimaa Ringed Seal Research Group of the University of Eastern Finland. Angel G. Rivera-Colón and Jouni Aspi provided guidance and helpful discussions on the population-genomic analyses. This study was supported by the US National Science Foundation (DEB-1239788, DEB-1342604, and DEB-1855812 to K.P.J) and the Academy of Finland (project 294466 to T.N.). Sequencing costs were supported by grants from the Oskar Öflund Foundation, the Betty Väänänen Foundation, Societas Pro Fauna et Flora Fennica, the Raija and Ossi Tuuliainen Foundation, and the Nestori Foundation.

References

Alcala N, Rosenberg NA. 2019. GST', Jost's D, and FST are similarly constrained by allele frequencies: a mathematical, simulation, and empirical study. *Mol. Ecol.* 28:1624–36

Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19:1655–64

516 Allen JM, Boyd B, Nguyen NP, Vachaspati P, Warnow T, et al. 2017. Phylogenomics from
517 whole genome sequences using aTRAM. *Syst. Biol.* 66:786–98

518 Allen JM, Huang DI, Cronk QC, Johnson KP. 2015. aTRAM - automated target restricted
519 assembly method: a fast method for assembling loci across divergent taxa from next-
520 generation sequencing data. *BMC Bioinformatics*. 16:e98

521 Appelgren ASC, Saladin V, Richner H, Doligez B, McCoy KD. 2018. Gene flow and adaptive
522 potential in a generalist ectoparasite. *BMC Evol. Biol.* 18:e99

523 Aydın M, Kryvoruchko IS, Şakiroğlu M. 2019. Widgetcon: a website and program for quick
524 conversion among common population genetic data formats. *Mol. Ecol. Resour.* 19:1374–
525 77

526 Baldwin-Brown JG, Villa SM, Vickrey AI, Johnson KP, Bush SE, et al. 2021. The assembled
527 and annotated genome of the pigeon louse Columbicola columbae, a model ectoparasite. *G3*
528 *Genes, Genomes, Genet.* 11:jkab009

529 Benesh DP, Weinreich F, Kalbe M, Milinski M. 2014. Lifetime inbreeding depression, purging,
530 and mating system evolution in a simultaneous hermaphrodite tapeworm. *Evolution (N. Y).*
531 68:1762–74

532 Berti E, Monsarrat S, Munk M, Jarvie S, Svenning JC. 2020. Body size is a good proxy for
533 vertebrate charisma. *Biol. Conserv.* 251:e108790

534 Biard V, Nykänen M, Niemi M, Kunnasranta M. 2022. Extreme moulting site fidelity of the
535 Saimaa ringed seal. *Mamm. Biol.* in press

536 Blasco-Costa I, Poulin R. 2013. Host traits explain the genetic structure of parasites: a meta-
537 analysis. *Parasitology*. 140:1316–22

538 Blomqvist D, Pauliny A, Larsson M, Flodin LA. 2010. Trapped in the extinction vortex? Strong
539 genetic effects in a declining vertebrate population. *BMC Evol. Biol.* 10:e33

540 Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. TrimAl: a tool for automated
541 alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*. 25:1972–73

542 Carlson CJ, Burgio KR, Dougherty ER, Phillips AJ, Bueno VM, et al. 2017. Parasite biodiversity
543 faces extinction and redistribution in a changing climate. *Sci. Adv.* 3:e1602422

544 Carlson CJ, Hopkins S, Bell KC, Doña J, Godfrey SS, et al. 2020. A global parasite conservation
545 plan. *Biol. Conserv.* 250:e108596

546 Cassinello J, Gomendio M, Roldan ERS. 2001. Relationship between coefficient of inbreeding
547 and parasite burden in endangered gazelles. *Conserv. Biol.* 15:1171–74

548 Cole R, Viney M. 2019. The population genetics of parasitic nematodes of wild animals. *Parasit.*
549 *Vectors*. 12:1–20

550 Coltman DW, Pilkington JG, Smith JA, Pemberton JM. 1999. Parasite-mediated selection
551 against inbred Soay sheep in a free-living, island population. *Evolution (N. Y).* 53:1259–67

552 Courchamp F, Clutton-Brock T, Grenfell B. 1999. Inverse density dependence and the Allee
553 effect. *Trends Ecol. Evol.* 14:405–10

554 Courchamp F, Jaric I, Albert C, Meinard Y, Ripple WJ, Chapron G. 2018. The paradoxical
555 extinction of the most charismatic animals. *PLoS Biol.* 16:e2003997

556 Danecek P, Auton A, Abecasis G, Albers CA, Banks E, et al. 2011. The variant call format and
557 VCFtools. *Bioinformatics*. 27:2156–58

558 de Moya RS, Yoshizawa K, Walden KKO, Sweet AD, Dietrich CH, Johnson KP. 2021.

559 Phylogenomics of parasitic and nonparasitic lice (Insecta: Psocodea): combining sequence
560 data and exploring compositional bias solutions in next generation data sets. *Syst. Biol.*
561 70:719–38

562 Detwiler JT, Criscione CD. 2017. Role of parasite transmission in promoting inbreeding: II.
563 Pedigree reconstruction reveals sib-transmission and consequent kin-mating. *Mol. Ecol.*
564 26:4405–17

565 DeWoody JA, Harder AM, Mathur S, Willoughby JR. 2021. The long-standing significance of
566 genetic diversity in conservation. *Mol. Ecol.* 30:4147–54

567 Dharmarajan G, Beasley JC, Beatty WS, Olson ZH, Fike JA, et al. 2016. Genetic co-structuring
568 in host-parasite systems: empirical data from raccoons and raccoon ticks. *Ecosphere*.
569 7:e01269

570 Dharmarajan G, Gupta P, Vishnudas CK, Robin V V. 2021. Anthropogenic disturbance favours
571 generalist over specialist parasites in bird communities: implications for risk of disease
572 emergence. *Ecol. Lett.* 24:1859–68

573 DiBlasi E, Johnson KP, Stringham SA, Hansen AN, Beach AB, et al. 2018. Phoretic dispersal
574 influences parasite population genetic structure. *Mol. Ecol.* 27:2770–79

575 Díez-del-Molino D, Sánchez-Barreiro F, Barnes I, Gilbert MTP, Dalén L. 2018. Quantifying
576 temporal genomic erosion in endangered species. *Trends Ecol. Evol.* 33:176–85

577 Doña J, Virrueta Herrera S, Nyman T, Kunnasranta M, Johnson KP. 2021. Patterns of
578 microbiome variation among infrapopulations of permanent bloodsucking parasites. *Front.*
579 *Microbiol.* 12:e642543

580 Dunn RR, Harris NC, Colwell RK, Koh LP, Sodhi NS. 2009. The sixth mass coextinction: are

most endangered species parasites and mutualists? *Proc. R. Soc. B Biol. Sci.* 276:3037–45

Ekroth AKE, Rafaluk-Mohr C, King KC. 2019. Host genetic diversity limits parasite success beyond agricultural systems: a meta-analysis. *Proc. R. Soc. B Biol. Sci.* 286:e20191811

Excoffier L, Lischer HELL. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10:564–67

Forsman A. 2014. Effects of genotypic and phenotypic variation on establishment are important for conservation, invasion, and infection biology. *Proc. Natl. Acad. Sci. USA*. 111:302–7

Fredericksen M, Ameline C, Krebs M, Hüssy B, Fields PD, et al. 2021. Infection phenotypes of a coevolving parasite are highly diverse, structured, and specific. *Evolution (N. Y).* 75:2540–54

Gagne RB, Crooks KR, Craft ME, Chiu ES, Fountain-Jones NM, et al. 2022. Parasites as conservation tools. *Conserv. Biol.* 36:e13719

Gómez A, Nichols E. 2013. Neglected wild life: parasitic biodiversity as a conservation target. *Int. J. Parasitol. Parasites Wildl.* 2:222–27

Gousy-Leblanc M, Yannic G, Therrien JF, Lecomte N. 2021. Mapping our knowledge on birds of prey population genetics. *Conserv. Genet.* 22:685–702

Harris NC, Livieri TM, Dunn RR. 2014. Ectoparasites in black-footed ferrets (Mustela nigripes) from the largest reintroduced population of the Conata Basin, South Dakota, USA. *J. Wildl. Dis.* 50:340–43

Haubold B, Pfaffelhuber P, Lynch M. 2010. MlRho - a program for estimating the population mutation and recombination rates from shotgun-sequenced diploid genomes. *Mol. Ecol.* 19(Suppl. 1):277–84

603 Hoffman JI, Simpson F, David P, Rijks JM, Kuiken T, et al. 2014. High-throughput sequencing
604 reveals inbreeding depression in a natural population. *Proc. Natl. Acad. Sci. USA*.
605 111:3775–80

606 Huyse T, Poulin R, Théron A. 2005. Speciation in parasites: a population genetics approach.
607 *Trends Parasitol.* 21:469–75

608 Itescu Y. 2019. Are island-like systems biologically similar to islands? A review of the evidence.
609 *Ecography (Cop.).* 42:1298–1314

610 IUCN. 2021. *The IUCN Red List of Threatened Species. Version 2021-1.*

611 Jakobsson M, Edge MD, Rosenberg NA. 2013. The relationship between FST and the frequency
612 of the most frequent allele. *Genetics*. 193:515–28

613 Johnson KP. 2019. Putting the genome in insect phylogenomics. *Curr. Opin. Insect Sci.* 36:111–
614 17

615 Johnson KP, Allen JM, Olds BP, Mugisha L, Reed DL, et al. 2014. Rates of genomic divergence
616 in humans, chimpanzees and their lice. *Proc. R. Soc. B Biol. Sci.* 281:13–17

617 Johnson KP, Walden KKO, Robertson HM. 2013. Next-generation phylogenomics using a
618 Target Restricted Assembly Method. *Mol. Phylogenet. Evol.* 66:417–22

619 Jombart T. 2008. Adegenet: a R package for the multivariate analysis of genetic markers.
620 *Bioinformatics*. 24:1403–5

621 Jombart T. 2016. *Introduction to genetic data analysis using R*. http://adegenet.r-forge.r-
622 project.org/files/PRstats/practical-MVAintro.1.0.pdf

623 Jombart T, Ahmed I. 2011. Adegenet 1.3-1: new tools for the analysis of genome-wide SNP

data. *Bioinformatics*. 27:3070–71

Kapusta A, Suh A, Feschotte C. 2017. Dynamics of genome size evolution in birds and mammals. *Proc. Natl. Acad. Sci. USA*. 114:E1460–69

Karamanlidis AA, Skrbinšek T, Amato G, Dendrinos P, Gaughran S, et al. 2021. Genetic and demographic history define a conservation strategy for earth's most endangered pinniped, the Mediterranean monk seal Monachus monachus. *Sci. Rep.* 11:e373

Kardos M, Taylor HR, Ellegren H, Luikart G, Allendorf FW. 2016. Genomics advances the study of inbreeding depression in the wild. *Evol. Appl.* 9:1205–18

Kim KC. 1975. Ecology and morphological adaptation of the sucking lice (Anoplura, Echinophthiriidae) on the northern fur seal. *Rapp. Procès verbaux des Réunions du Cons. Perm. Int. pour l'Exploration la Mer*. 169:504–15

Koop JAH, DeMatteo KE, Parker PG, Whiteman NK. 2014. Birds are islands for parasites. *Biol. Lett.* 10:e20140255

Kunnasranta M, Niemi M, Auttila M, Valtonen M, Kammonen J, Nyman T. 2021. Sealed in a lake — biology and conservation of the endangered Saimaa ringed seal: a review. *Biol. Conserv.* 253:e108908

Kwak ML, Heath ACG, Cardoso P. 2020. Methods for the assessment and conservation of threatened animal parasites. *Biol. Conserv.* 248:e108696

Kyriazis CC, Wayne RK, Lohmueller KE. 2021. Strongly deleterious mutations are a primary determinant of extinction risk due to inbreeding depression. *Evol. Lett.* 5:33–47

Lande R. 1993. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *Am. Nat.* 142:911–27

646 Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods*.
647 9:357–59

648 Leblois R, Estoup A, Rousset F. 2003. Influence of mutational and sampling factors on the
649 estimation of demographic parameters in a "continuous" population under isolation by
650 distance. *Mol. Biol. Evol.* 20:491–502

651 Legendre P. 2018. *Package "lmodel2."* https://cran.r-
652 project.org/web/packages/lmodel2/lmodel2.pdf

653 Leidenberger S, Harding K, Härkönen T. 2007. Phocid seals, seal lice and heartworms: a
654 terrestrial host-parasite system conveyed to the marine environment. *Dis. Aquat. Organ.*
655 77:235–53

656 Lemopoulos A, Prokkola JM, Uusi-Heikkilä S, Vasemägi A, Huusko A, et al. 2019. Comparing
657 RADseq and microsatellites for estimating genetic diversity and relatedness — Implications
658 for brown trout conservation. *Ecol. Evol.* 9:2106–20

659 Leo NP, Hughes JM, Yang X, Poudel SKS, Brogdon WG, Barker SC. 2005. The head and body
660 lice of humans are genetically distinct (Insecta: Phthiraptera, Pediculidae): evidence from
661 double infestations. *Heredity (Edinb).* 95:34–40

662 Leonardi MS, Crespo EA, Raga JA, Aznar FJ. 2013. Lousy mums: patterns of vertical
663 transmission of an amphibious louse. *Parasitol. Res.* 112:3315–23

664 Leonardi MS, Virrueta Herrera S, Sweet A, Negrete J, Johnson KP. 2019. Phylogenomic
665 analysis of seal lice reveals codivergence with their hosts. *Syst. Entomol.* 44:699–708

666 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. 2009. The Sequence Alignment/Map
667 format and SAMtools. *Bioinformatics*. 25:2078–79

668 Loiselle BA, Sork VL, Nason J, Graham C. 1995. Spatial genetic structure of a tropical
669 understory shrub, Psychotria officinalis (Rubiaceae). *Am. J. Bot.* 82:1420–25

670 Luo SJ, Liu YC, Xu X. 2019. Tigers of the world: genomics and conservation. *Annu. Rev. Anim.*
671 *Biosci.* 7:521–48

672 Martinez-Bakker ME, Sell SK, Swanson BJ, Kelly BP, Tallmon DA. 2013. Combined genetic
673 and telemetry data reveal high rates of gene flow, migration, and long-distance dispersal
674 potential in Arctic ringed seals (Pusa hispida). *PLoS One*. 8:e77125

675 Martinů J, Štefka J, Poosakkannu A, Hypša V. 2020. "Parasite turnover zone" at secondary
676 contact: a new pattern in host–parasite population genetics. *Mol. Ecol.* 29:4653–64

677 Mazé-Guilmo E, Blanchet S, Mccoy KD, Loot G. 2016. Host dispersal as the driver of parasite
678 genetic structure: a paradigm lost? *Ecol. Lett.* 19:336–47

679 McCoy KD, Boulinier T, Tirard C. 2005. Comparative host-parasite population structures:
680 disentangling prospecting and dispersal in the black-legged kittiwake Rissa tridactyla. *Mol.*
681 *Ecol.* 14:2825–38

682 Meirmans PG. 2020. Genodive version 3.0: easy-to-use software for the analysis of genetic data
683 of diploids and polyploids. *Mol. Ecol. Resour.* 20:1126–31

684 Meirmans PG, Hedrick PW. 2011. Assessing population structure: FST and related measures.
685 *Mol. Ecol. Resour.* 11:5–18

686 Meyer M, Kircher M, Gansauge MT, Li H, Racimo F, et al. 2012. A high-coverage genome
687 sequence from an archaic Denisovan individual. *Science (80-.).* 338:222–26

688 Milotic M, Lymbery A, Thompson A, Doherty JF, Godfrey S. 2020. Parasites are endangered by
689 the conservation of their hosts: meta-analyses of the effect of host captivity on the odds of

parasite infection. *Biol. Conserv.* 248:e108702

Mirarab S, Nguyen N, Guo S, Wang L-S, Kim J, Warnow T. 2015. PASTA: ultra-large multiple sequence alignment for nucleotide and amino-acid sequences. *J. Comput. Biol.* 22:377–86

Mirarab S, Reaz R, Bayzid MS, Zimmermann T, S. Swenson M, Warnow T. 2014. ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics*. 30:541–48

Nagel R, Stainfield C, Fox-Clarke C, Toscani C, Forcada J, Hoffman JI. 2021. Evidence for an Allee effect in a declining fur seal population. *Proc. R. Soc. B Biol. Sci.* 288:e20202882

Nyman T, Valtonen M, Aspi J, Ruokonen M, Kunnasranta M, Palo JU. 2014. Demographic histories and genetic diversities of Fennoscandian marine and landlocked ringed seal subspecies. *Ecol. Evol.* 4:3420–34

O'Brien SJ, Johnson WE, Driscoll CA, Dobrynin P, Marker L. 2017. Conservation genetics of the cheetah: lessons learned and new opportunities. *J. Hered.* 108:671–77

Orsini L, Vanoverbeke J, Swillen I, Mergeay J, De Meester L. 2013. Drivers of population genetic differentiation in the wild: isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Mol. Ecol.* 22:5983–99

Palo JU, Hyvärinen H, Helle E, Mäkinen HS, Väinölä R. 2003. Postglacial loss of microsatellite variation in the landlocked Lake Saimaa ringed seal. *Conserv. Genet.* 4:117–28

Palo JU, Mäkinen HS, Helle E, Stenman O, Väinölä R. 2001. Microsatellite variation in ringed seals (Phoca hispida): genetic structure and history of the Baltic Sea population. *Heredity (Edinb).* 86:609–17

Peart CR, Tusso S, Pophaly SD, Botero-Castro F, Wu CC, et al. 2020. Determinants of genetic variation across eco-evolutionary scales in pinnipeds. *Nat. Ecol. Evol.* 4:1095–1104

712 Pérez JM, Meneguz PG, Dematteis A, Rossi L, Serrano E. 2006. Parasites and conservation
713 biology: the "ibex-ecosystem." *Biodivers. Conserv.* 15:2033–47

714 Pérez JM, Sánchez I, Palma RL. 2013. The dilemma of conserving parasites: the case of Felicola
715 (Lorisicola) isidoroi (Phthiraptera: Trichodectidae) and its host, the endangered Iberian lynx
716 (Lynx pardinus). *Insect Conserv. Divers.* 6:680–86

717 Price PW. 1980. *Evolutionary Biology of Parasites*. Princeton, NJ: Princeton University Press

718 R_Core_Team. 2021. *R: a language and environment for statistical computing*. https://www.r-
719 project.org/

720 Rey-Iglesia A, Gaubert P, Espregueira Themudo G, Pires R, De La Fuente C, et al. 2021.
721 Mitogenomics of the endangered Mediterranean monk seal (Monachus monachus) reveals
722 dramatic loss of diversity and supports historical gene-flow between Atlantic and eastern
723 Mediterranean populations. *Zool. J. Linn. Soc.* 191:1147–59

724 Rochette NC, Rivera-Colón AG, Catchen JM. 2019. Stacks 2: Analytical methods for paired-end
725 sequencing improve RADseq-based population genomics. *Mol. Ecol.* 28:4737–54

726 Rousset F. 2000. Genetic differentiation between individuals. *J. Evol. Biol.* 13:58–62

727 Rousset F. 2008. GENEPOP'007: a complete re-implementation of the GENEPOP software for
728 Windows and Linux. *Mol. Ecol. Resour.* 8:103–6

729 Rózsa L, Vas Z. 2015. Co-extinct and critically co-endangered species of parasitic lice, and
730 conservation-induced extinction: should lice be reintroduced to their hosts? *Oryx*. 49:107–
731 10

732 Sayyari E, Mirarab S. 2016. Fast coalescent-based computation of local branch support from
733 quartet frequencies. *Mol. Biol. Evol.* 33:1654–68

734 Solórzano-García B, Vázquez-Domínguez E, Pérez-Ponce de León G, Piñero D. 2021. Co-
735 structure analysis and genetic associations reveal insights into pinworms (Trypanoxyuris)
736 and primates (Alouatta palliata) microevolutionary dynamics. *BMC Ecol. Evol.* 21:e190

737 Spencer HG, Zuk M. 2016. For host's sake: the pluses of parasite preservation. *Trends Ecol.*
738 *Evol.* 31:341–43

739 Spielman D, Brook BW, Frankham R. 2004. Most species are not driven to extinction before
740 genetic factors impact them. *Proc. Natl. Acad. Sci. USA*. 101:15261–64

741 Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
742 large phylogenies. *Bioinformatics*. 30:1312–13

743 Stoffel MA, Humble E, Paijmans AJ, Acevedo-Whitehouse K, Chilvers BL, et al. 2018.
744 Demographic histories and genetic diversity across pinnipeds are shaped by human
745 exploitation, ecology and life-history. *Nat. Commun.* 9:e4836

746 Stork NE, Lyal CHC. 1993. Extinction or "co-extinction" rates? *Nature*. 366:307

747 Strona G. 2015. Past, present and future of host-parasite co-extinctions. *Int. J. Parasitol.*
748 *Parasites Wildl.* 4:431–41

749 Sunde J, Yıldırım Y, Tibblin P, Forsman A. 2020. Comparing the performance of microsatellites
750 and RADseq in population genetic studies: analysis of data for pike (Esox lucius) and a
751 synthesis of previous studies. *Front. Genet.* 11:e218

752 Sweet AD, Johnson KP. 2018. The role of parasite dispersal in shaping a host–parasite system at
753 multiple evolutionary scales. *Mol. Ecol.* 27:5104–19

754 Sweet AD, Wilson RE, Sonsthagen SA, Johnson KP. 2020. Lousy grouse: comparing
755 evolutionary patterns in Alaska galliform lice to understand host evolution and host–

parasite interactions. *Ecol. Evol.* 10:8379–93

Taavitsainen JP. 2005. Finland in the Middle Ages. *Archaeol. J.* 162(Supplement 1):11–17

Thompson BS, Rog SM. 2019. Beyond ecosystem services: using charismatic megafauna as flagship species for mangrove forest conservation. *Environ. Sci. Policy*. 102:9–17

Thompson RCA, Lymbery AJ, Godfrey SS. 2018. Parasites at risk – insights from an endangered marsupial. *Trends Parasitol.* 34:12–22

Valtonen M, Heino M, Aspi J, Buuri H, Kokkonen T, et al. 2015. Genetic monitoring of a critically-endangered seal population based on field-collected placentas. *Ann. Zool. Fennici*. 52:51–65

Valtonen M, Palo JUU, Aspi J, Ruokonen M, Kunnasranta M, Nyman T. 2014. Causes and consequences of fine-scale population structure in a critically endangered freshwater seal. *BMC Ecol.* 14:e22

Valtonen M, Palo JU, Ruokonen M, Kunnasranta M, Nyman T. 2012. Spatial and temporal variation in genetic diversity of an endangered freshwater seal. *Conserv. Genet.* 13:1231–45

Van Den Broeck F, Meurs L, Raeymaekers JAM, Boon N, Dieye TN, et al. 2014. Inbreeding within human Schistosoma mansoni: do host-specific factors shape the genetic composition of parasite populations? *Heredity (Edinb).* 113:32–41

[dataset]Virrueta Herrera S, Johnson KP, Sweet AD, Ylinen E, Kunnasranta M, Nyman T. 2022. *Data for: High levels of inbreeding with spatial and host-associated structure in lice of an endangered freshwater seal*. Dryad Digital Repository. https://doi.org/10.5061/dryad.8sf7m0cqr

Vlasman K, Campbell D. 2004. *Field Guide: Diseases and Parasites of Marine Mammals of the*

778 *Eastern Arctic*. Guelph: Canadian Cooperative Wildlife Health Centre

779 Whiteman NK, Parker PG. 2005. Using parasites to infer host population history: a new rationale
780 for parasite conservation. *Anim. Conserv.* 8:175–81

781 Williams NF, McRae L, Freeman R, Capdevila P, Clements CF. 2021. Scaling the extinction
782 vortex: body size as a predictor of population dynamics close to extinction events. *Ecol.*
783 *Evol.* 11:7069–79

784 Windsor DA. 1997. Equal rights for parasites. *Perspect. Biol. Med.* 49:222–29

785 Zarowiecki M, Berriman M. 2015. What helminth genomes have taught us about parasite
786 evolution. *Parasitology*. 142:S85–97

Data accessibility and benefit-sharing section

Data accessibility statement

Raw sequence reads are deposited in the SRA, under the SRA and BioProject accession numbers listed in Supporting Table S1, which also contains the metadata for each louse individual. Data files used in the phylogenomic and population-genomic analyses have been deposited in the Dryad repository (link for review: https://datadryad.org/stash/share/62coO0qXpHGopyud1DF2_rgoNtFfzJJHtRzyFdRFvQ8).

Benefit-sharing statement

Benefits from this research accrue from presenting information on the biology of an endemic and endangered host–parasite system and the sharing of our data and results on public databases as described above. This study complies with laws governing handling of endangered animals (see Methods).

Author contributions

K.P.J., T.N. and S.V.H. conceived the study. M.K. and E.Y. obtained samples. S.V.H. and K.P.J. collected the data. S.V.H, A.D.S., and T.N. analyzed the data. T.N., M.K., E.Y., and K.P.J. obtained financial support for the project. T.N. and S.V.H. wrote the manuscript, and all authors contributed to editing the manuscript.

Tables

Table 1. Results of the hierarchical locus-by-locus AMOVA when individual lice are grouped according to three main lake areas (Fig. 2A) and host seal individuals (infrapopulations) within the areas. The effect of all explanatory variables is significantly different from 0 at $P < 0.0001$.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among lake areas	(2)	2653.232	40.78219	12.99
Among infrapopulations within lake areas	(15)	6452.108	67.27029	21.43
Between louse individuals within infrapopulations	(18)	2890.500	-44.45757	-14.16
Within louse individuals	(36)	9002.000	250.26769	79.74
Total		20997.840	313.86260	

Figure legends

Fig. 1. (A) *Echinophthirius horridus* seal louse male (top) and female (bottom) from Lake Saimaa (for both, ventral view on left and dorsal on right). (B) Seal lice on the muzzle of a dead Saimaa ringed seal; the white arrow shows one of three individuals. (C) Saimaa ringed seal female nursing a weaning-age pup.

Fig. 2. (A) Map Lake Saimaa, with collection sites of seals and their paired lice shown by colored dots that are labeled with the seal and louse individual numbers. The main basins of the Lake Saimaa complex are separated by broken lines, with area names indicated on the side. The location of the town Savonlinna at the Kyrönsalmi strait is indicated by a red circle. (B) PCoA ordination plot of seal lice based on their genetic similarity. Lice from the same seal are colored similarly and connected by lines, and dot colors and shadings correspond to those used in panel A. Dot shading indicates the main lake area (see legend). Note that lice from Northern Saimaa and Haukivesi tend to be located on the right-hand side of the ordination, while lice from the two southernmost areas are to the left. (C) Relationship between genetic distance and ln geographic distance between individual seal lice in the full dataset. (D) Admixture plot for individual seal lice at $K = 2$. Section heights within bars show the proportion of ancestry attributed to "northern" (blue) and "southern" (orange) ancestry. Louse individuals are denoted below the plot and ordered from the south to north in the left to right direction, the main lake areas are indicated above the plot, and the locations of the borders between them (see panel A) are indicated by inverted triangles. The location of the town Savonlinna at the Kyrönsalmi strait is indicated by a red circle above the triangle.

Fig. 3. Correlation between mean θ estimates of lice collected from the same seal individuals (*i.e.*, same infrapopulation). Dot colors correspond to those used in Fig. 1A, labels indicate the

seal individual from which the lice were collected, and dot shading shows the lake area (see legend). The red line represents the correlation from a reduced major axis regression, and grey lines represent the confidence limits of the slope.

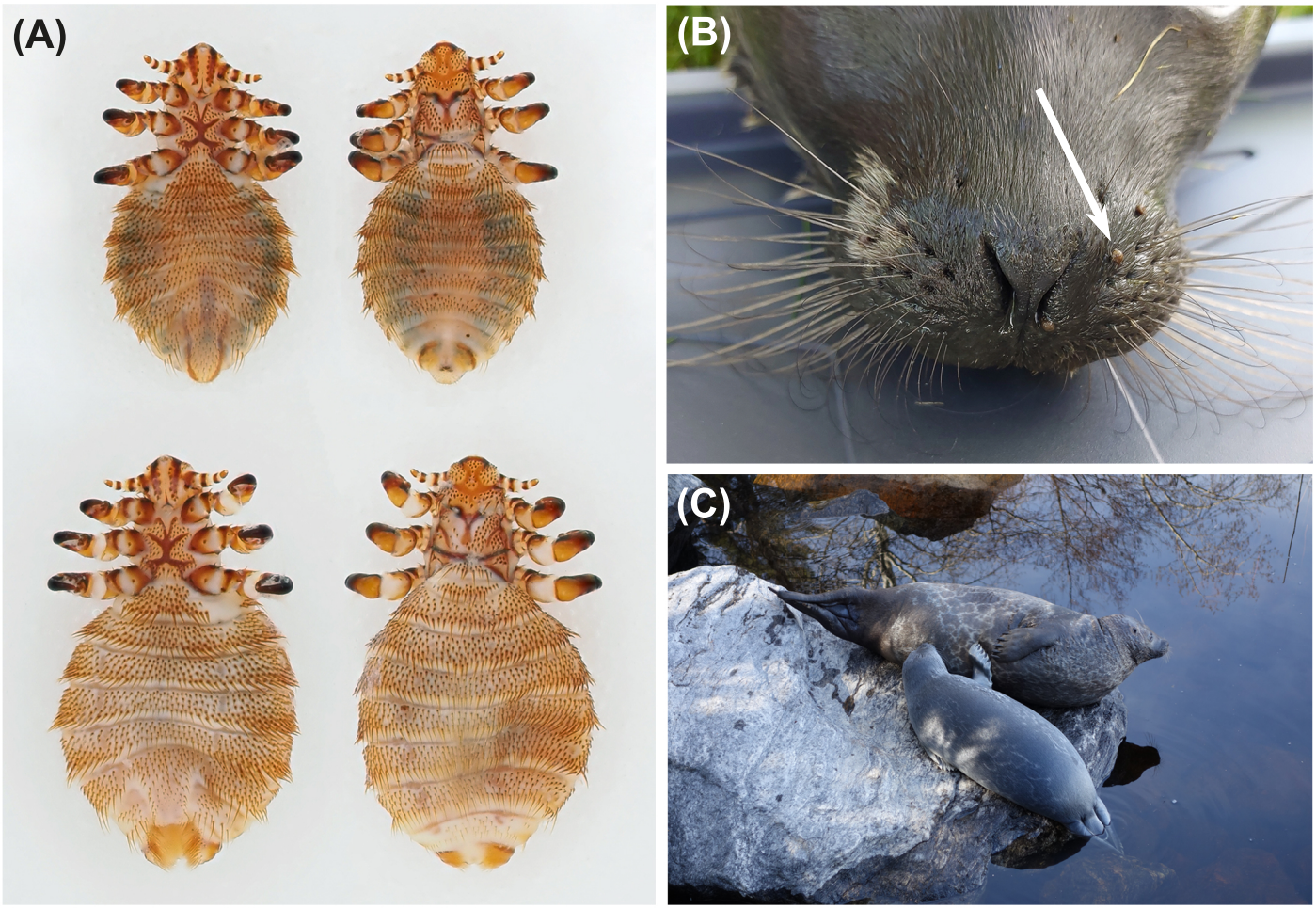

Fig. 1

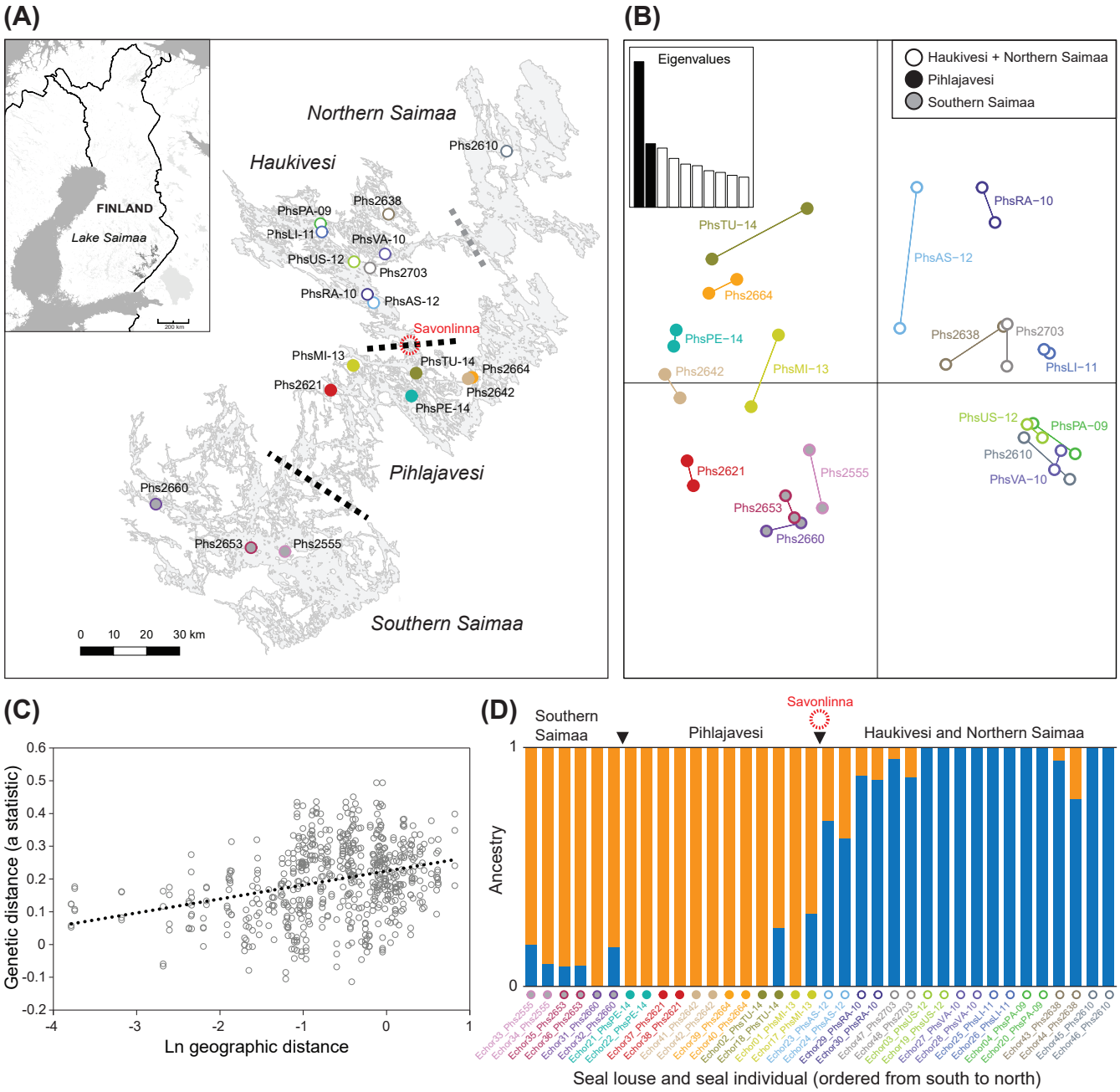
(A)
FINLAND
Lake Saimaa
Northern Saimaa
Haukivesi
Pihlajavesi
Southern Saimaa
Savonlinna
Phs2610
Phs2638
PhsPA-09
PhsLI-11
PhsVA-10
PhsUS-12
Phs2703
PhsRA-10
PhsAS-12
PhsMI-13
PhsTU-14
Phs2664
Phs2621
Phs2642
PhsPE-14
Phs2660
Phs2653
Phs2555
0 10 20 30 km
(B)
Eigenvalues
Haukivesi + Northern Saimaa
Pihlajavesi
Southern Saimaa
PhsTU−14
Phs2664
PhsPE−14
Phs2642
PhsMI−13
Phs2621
Phs2555
Phs2653
Phs2660
PhsRA−10
PhsAS−12
Phs2638
Phs2703
PhsLI−11
PhsUS−12
PhsPA−09
Phs2610
PhsVA−10
(C)
Genetic distance (a statistic)
Ln geographic distance
(D)
Southern Saimaa
Pihlajavesi
Savonlinna
Haukivesi and Northern Saimaa
Ancestry
Seal louse and seal individual (ordered from south to north)

Fig. 2

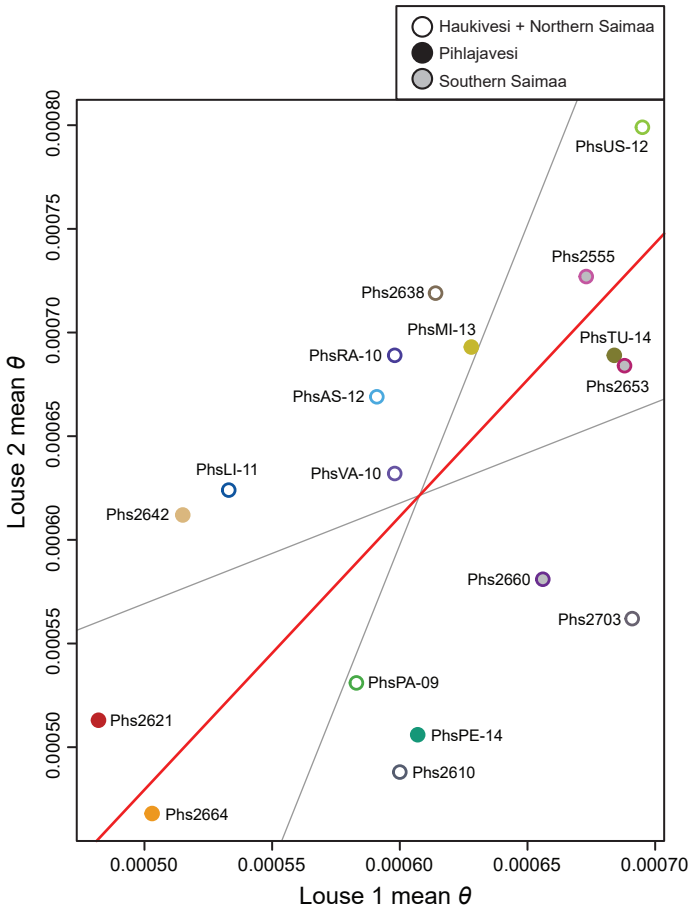
Haukivesi + Northern Saimaa
Pihlajavesi
Southern Saimaa
PhsUS-12
Phs2555
Phs2638
PhsMI-13
PhsTU-14
PhsRA-10
Phs2653
PhsAS-12
PhsLI-11
PhsVA-10
Phs2642
Phs2660
Phs2703
PhsPA-09
Phs2621
PhsPE-14
Phs2610
Phs2664
0.00050
0.00055
0.00060
0.00065
0.00070
0.00075
0.00080
Louse 1 mean θ
Louse 2 mean θ

Fig. 3