Evolutionary history and seascape genomics of Harbour porpoises (Phocoena phocoena) across environmental gradients in the North Atlantic and adjacent waters

Enrique Celemín[1], Marijke Autenrieth[2], Anna Roos[1], Iwona Pawliczka[1], María Quintela[3], Ulf Lindstrøm[1], Harald Benke[1], Ursula Siebert[4], Christina Lockyer[1], Per Berggren[1], Ayaka Öztürk[1], Bayram Öztürk[5], Véronique Lesage[1], and Ralph Tiedemann[1]

[1]Affiliation not available
[2]Universitat Potsdam
[3]Havforskningsinstituttet
[4]University of Veterinary Medicine Hannover, Institute for Terrestrial and Aquatic Wildlife Research
[5]Faculty of Fisheries, Istanbul University

December 21, 2022

Abstract

The Harbour porpoise (Phocoena phocoena) is a highly mobile cetacean species which primarily occurs in coastal and shelf waters across the Northern hemisphere. It inhabits heterogeneous seascapes that vary broadly in salinity and temperature. Here we produced 74 whole genomes at intermediate coverage to study Harbour porpoise's evolutionary history and investigate the role of local adaptation in the diversification into subspecies and populations. We identified ~6 million high quality SNPs sampled at 8 localities across the North Atlantic and adjacent waters, which we used for population structure, demographic, and genotype-environment association analyses. Our results support a genetic differentiation between three subspecies, and three distinct populations within the subspecies P.p. phocoena: Atlantic, Belt Sea and Proper Baltic Sea. Effective population size and Tajima's D levels suggest a population contraction in both Black Sea and Iberian porpoises while a population expansion in the P.p. phocoena populations. Phylogenetic trees indicate a post-glacial colonization of Harbour porpoises from a southern refugium. Genotype-environment association analysis identified salinity as a major driver in genomic variation and we identified candidate genes putatively underlying adaptation to different salinity levels. Our study highlights the value of whole genome resequencing to unravel subtle population structure in highly mobile species and shows how strong environmental gradients and local adaptation may lead to population differentiation. The results have great conservation implications as we found major levels of inbreeding and low genetic diversity in the endangered Black Sea subspecies and identified the critically endangered Proper Baltic Sea porpoises as a separate population.

Evolutionary history and seascape genomics of Harbour porpoises (*Phocoena phocoena*) across environmental gradients in the North Atlantic and adjacent waters

Enrique Celemín[1], Marijke Autenrieth[1], Anna Roos[2], Iwona Pawliczka[3], María Quintela[4], Ulf Lindstrøm[5,6], Harald Benke[7], Ursula Siebert[8], Christina Lockyer[9], Per Berggren[10], A. Amaha Özturk[11], Bayram Özturk[11], Véronique Lesage[12], Ralph Tiedemann[1,*]

[1] University of Potsdam, Institute of Biochemistry and Biology, Evolutionary Biology & Systematic Zoology, Potsdam, Germany

[2] Swedish Museum of Natural History, Stockholm, Sweden

[3] Prof. Krzysztof Skóra Hel Marine Station, Faculty of Oceanography and Geography, University of Gdansk, Hel, Poland

[4] Institute of Marine Research, Bergen, Norway

[5] Institute of Marine Research, Tromsø, Norway

[6] UiT Norwegian Arctic University, Tromsø, Norway

[7] Deutsches Meeresmuseum, Stralsund, Germany

[8] Institute for Terrestrial and Aquatic Wildlife Research (ITAW), University of Veterinary Medicine Hannover Foundation, Büsum, Germany

[9] Age Dynamics, 2800 Kongens Lyngby, Denmark

[10] School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, United Kingdom

[11] Marine Biology Department, Faculty of Fisheries, Istanbul University, Istanbul, Turkey

[12] Fisheries and Oceans Canada, Maurice Lamontagne Institute, Mont-Joli, Quebec, Canada

* Corresponding author; email: tiedeman@uni-potsdam.de

Abstract

The Harbour porpoise (*Phocoena phocoena*) is a highly mobile cetacean species which primarily occurs in coastal and shelf waters across the Northern hemisphere. It inhabits heterogeneous seascapes that vary broadly in salinity and temperature. Here we produced 74 whole genomes at intermediate coverage to study Harbour porpoise's evolutionary history and investigate the role of local adaptation in the diversification into subspecies and populations. We identified ~6 million high quality SNPs sampled at 8 localities across the North Atlantic and adjacent waters, which we used for population structure, demographic, and genotype-environment association analyses. Our results support a genetic differentiation between three subspecies, and three distinct populations within the subspecies*P.p. phocoena* : Atlantic, Belt Sea and Proper Baltic Sea. Effective population size and Tajima's D levels suggest a population contraction in both Black Sea and Iberian porpoises while a population expansion in the *P.p. phocoena* populations. Phylogenetic trees indicate a post-glacial colonization of Harbour porpoises from a southern refugium. Genotype-environment association analysis identified salinity as a major driver in genomic variation and we identified candidate genes putatively underlying adaptation to different salinity levels. Our study highlights the value of whole genome resequencing to unravel subtle population structure in highly mobile species and shows how strong environmental gradients and local adaptation may lead to population differentiation. The results have great conservation implications as we found major levels of inbreeding and low genetic diversity in the endangered Black Sea subspecies and identified the critically endangered Proper Baltic Sea porpoises as a separate population.

Keywords: Harbour Porpoise, Genomics, Genetic Structure, Local Adaptation, Conservation

1. Introduction

The wellbeing of populations depends on an array of intrinsic and extrinsic factors, but effective population size and genetic diversity are among the most crucial. Reduction of genetic diversity in a population can lead to detrimental outcomes such as loss of adaptive potential and inbreeding depression, produced by the accumulation of slightly deleterious mutations due to reduced efficiency of purifying selection (Tanaka, 2000). Understanding the processes that influence current genetic variation is essential for the management and conservation of the diversity of a species (Kardos et al., 2021). Variation in genetic diversity is the result of

historical and present demographic, geographic, ecological and behavioral mechanisms that influence gene flow, genetic drift and/or selection levels (Stange, Barret & Hendry, 2021). Overarching processes such as glacial contractions and post-glacial expansions have influenced current patterns of genetic structure and diversity in many European species, creating population subdivisions and hybrid zones by secondary contact (Hewitt, 1999; 2000; 2001). However, neutral evolutionary processes are not the only factor contributing to population subdivisions, as selective processes such as local adaptation could also produce different evolutionary trajectories (Barret & Schluter, 2008). Therefore, both neutral and adaptive processes must be considered when studying genetic diversity and population structure. Dispersal ability over vast geographical distances may facilitate gene flow among distant locations and hence hinder population differentiation (Slatkin, 1987). The marine habitat is a perfect example of an environment where a lack of physical barriers offers a continuous environment such that highly mobile species could present large homogenous populations and an absence of genetic structure. Yet, several cetacean species show fine-scale population structure, as for instance Northern bottlenose whale (de Greef et al., 2022), Finless porpoise (Zhou et al., 2018), Killer whale (Foote et al., 2016), and bottlenose dolphins (Louis et al., 2021).

The Harbour porpoise (*Phocoena phocoena*) is a great example of complex genetic differentiation in a highly mobile cetacean species. It belongs to the *Phocoenidae* family, a group of seven species (Ben Chehida et al., 2020) containing some of the most threatened cetaceans (Carlén, Nunny & Simmonds, 2021): the Vaquita (*Phocoena sinus*), one of the most endangered mammals on planet earth, with a population size of ˜10 individuals (Jaramillo-Legorreta et al., 2019), and the Yangtze finless porpoise (*Neophocaena asiaeorientalis asiaeorientalis),* with only ˜1,000 individuals extant (Zhao et al., 2008). Porpoises are mainly affected by incidental bycatch (Brownell et al., 2019), pollutants such as PCBs (Berggren et al., 1999; Karlson et al., 2000), parasites (Dzido et al., 2021; Reckendorf et al., 2021; Ryeng et al., 2022), and noise pollution issued from offshore infrastructure developments, shipping routes and underwater explosions (Siebert et al., 2022). Harbour porpoises inhabit coastal and shelf waters across the Northern hemisphere and at least three subspecies have been described:*P.p. vomeria* in the North Pacific, *P.p. phocoena* in the North Atlantic, and *P.p. relicta* in the Black Sea. A fourth subspecies near the Iberian Peninsula and in Mauritanian waters (*P.p. meridionalis*) has been proposed (Fontaine et al., 2007; 2014), although a formal description has not yet been made.

The North Atlantic subspecies (*P.p. phocoena*) has a continuous distribution in the North Atlantic extending from the French Biscayan waters to the Baltic and Barents Sea and from the Norwegian Sea to the western North Atlantic coast of Canada and the United States, crossing Faroese, Icelandic and Greenlandic waters (Gaskin, 1992; Read 1999). The genetic structure of North Atlantic harbour porpoises has been widely studied (Alfonsi et al., 2012; Luna et al., 2012; Quintela et al., 2020 among others) and microsatellite data suggests that both sides of the Atlantic belong to the same lineage (Chehida et al., 2021). However, genetic and distribution densities suggest the presence of different ecotypes or populations on the peripheral waters of the Baltic Sea (Tiedemann et al., 1996; Wiemann et al., 2010; Lah et al., 2016) and West Greenland (Olsen et al., 2022). The Baltic Sea is a remarkable sub-basin of the Atlantic Ocean formed less than 10,000 years before present (BP) as a postglacial marine environment. Baltic populations of several marine organisms are genetically distinct from conspecifics from the North Sea and the Atlantic, possibly due to isolation, bottlenecks, and local adaptation (Wennerstrom et al., 2013). A series of small basins are separated by shallow underwater ridges ranging from the North Sea through Skagerrak, Kattegat, Belt Seas to the entrance of the proper Baltic Sea, making dispersal and gene flow limited (Johannesson & André, 2006). Moreover, the Baltic Sea is an extreme marine environment with low winter temperatures and one of the strongest salinity gradients in the world, ranging from ˜34 practical salinity units (psu) in the Skagerrak to ˜2 psu in the innermost parts of the Baltic (Feistel et al., 2010). These conditions make Baltic species a prime system to study local adaptation (DeFaveri et al., 2014; Wrange et al., 2014; Sjöqvist et al., 2015) and speciation in the marine environment (Riginos & Cunningham, 2005, Stuckas et al., 2009; Pereyra et al., 2009). Passive acoustic (Carlén et al., 2018), telemetry (Svegaard et al., 2015), morphological (Huggenberger et al., 2002; Galatius et al., 2012), and genetic data (Tiedemann et al., 1996; Wiemann et al., 2010; Lah et al., 2016) suggest the presence of three harbour porpoise populations in the Baltic region: one in the North

Sea, Skagerrak and northern parts of Kattegat (North Sea population, NOS), another in southern parts of Kattegat and Belt Seas (Belt Sea population, BES) and a third one in the Baltic Proper (Proper Baltic Sea population, PBS). Although overlap between the three populations has been reported based on both genetic (Wiemann et al., 2010; Lah et al., 2016) and satellite tracking data (Svegaard et al., 2015), borders among them have been postulated based on geographical separation during the reproductive season (Svegaard et al., 2015; Carlén et al., 2018, Amundin et al., 2022).

Harbour porpoise abundance estimates vary greatly among regions: Black Sea porpoise population size is unknown, but declined by ~90% between the 1930s and the 1980s (Birkun, 2002); the European Atlantic Shelf is estimated to be inhabited by ~375,000 individuals (Hammond et al., 2013), ~20,000 animals are estimated in the Belt Sea and only ~500 in the Proper Baltic Sea (Amundin et al., 2022). The Black Sea subspecies and the Proper Baltic Sea population are considered endangered and critically endangered, respectively. To date, no assessment of genetic diversity, N_e or population structure has been conducted on North Atlantic porpoises at the whole-genome level. Historically, conservation and evolutionary geneticists have leaned on a handful of molecular markers, such as mitochondrial DNA and microsatellites, for the study of genetic variation among populations (Schweizer et al., 2021). However, with the development of high-throughput sequencing there has been a transition from genetics to genomics (Formenti et al., 2022). The ever-decreasing cost of reduced-representation and whole genome sequencing methods has positioned conservation genomics as a prominent tool for the characterization of biodiversity and preservation of species (Fuentes-Pardo & Ruzzante, 2017). Nowadays, the democratization of sequencing costs allows to resequence the genome of a set of individuals to assess genetic variation across thousands or millions of markers and address long-standing questions in evolutionary biology not fully resolved with traditional markers or reduced-representation methods (Foote et al., 2021; Robinson et al., 2022; Wolf et al., 2022). This increase of statistical power to unravel subtle patterns not fully captured by less dense datasets has had and will continue to have a remarkable impact in the field of conservation genomics (Lou et al., 2021; Szarmach et al., 2021).

Here, we used genomics approaches to study the population structure, genetic diversity, evolutionary history and local adaptation of the Harbour porpoise (*Phocoena phocoena*). We generated the most comprehensive data set of North Atlantic harbour porpoises so far, by resequencing the whole-genome of 74 harbour porpoises from eight regions across the North Atlantic and adjacent waters. Our results shed light on the expansion of harbour porpoise populations across the North Atlantic, demonstrate that genome-wide data can unravel subtle population structure and contribute to understand how marine species adapt to their local environment. The results have great conservation implications as we found major levels of inbreeding and low genetic diversity in the endangered Black Sea subspecies and identified the critically endangered Proper Baltic Sea porpoises as a separate population.

2. Material and Methods

2.1 Sampling and laboratory procedures

The study was based on 74 tissue samples collected in eight different regions across the North Atlantic and adjacent waters (Figure 1): eastern Canada (CA), Iceland (ICE), Barents Sea (BAS), North Sea (NOS), Belt Sea (BES), Proper Baltic Sea (PBS), Iberia (IBE), and Black Sea (BLS). The NOS-BES border was located at the latitude 56.95^{o}N, as a straight line from Denmark to Sweden (Sveegaard et al., 2015), while the BES-PBS border was placed as a diagonal line from the Swedish Hanö island (56^{o}N 14.7^{o}E) to the village of Jarosławiec in Poland (54.5^{o}N 16.5^{o}E) (Carlén et al 2018; Amundin et al., 2022). All sampling was performed on by-caught or stranded carcasses, and no live harbour porpoise has been targeted in the scope of this study.

We extracted total genomic DNA from skin or muscle tissue using one of the three following methods: NucleoSpin Tissue Kit, DNeasy Blood & Tissue Kit, or Phenol-Chloroform extraction. DNA concentration and quality were measured using a Qubit Fluorometer and Fragment Analyzer to ensure that chosen samples were not fragmented and at least 300ng of DNA per sample were available. Library and whole-genome resequencing was performed at GENEWIZ from Azenta Life Sciences in Leipzig, Germany, on a NovaSeq

6000 platform with a S4 flow cell and using 150-bp paired-end reads.

2.2 Genomic data processing

Raw reads were demultiplexed and the individual *fastq* files were processed with the software *Fastp* (v.023.2) (Chen et al., 2018) to trim residual adapter sequences and poly G tails as well as to filter out bad/low quality (<15Q) and too short (<75bp) reads. The remaining filtered reads were mapped to the harbour porpoise reference genome assembly (Autenrieth et al., 2018) using the *Bwa mem* algorithm (v.0.7.17) (Li & Durbin, 2009) with default settings. Alignment *sam* files were converted to *bam* files and sorted by its leftmost coordinate with *Samtools (v.1.15)* (Danecek et al., 2021). *Picard tools* (v.2.27.2) was used to add read groups and to remove PCR and sequencing duplicates. Thereafter, *bam*files were realigned around indels with *GATK* (v.3.8.1) (Van der Auwera et al., 2013).

We used *RepeatMasker* (v4.1.2) (Smit et al., 2013) and the*dfam 3.6* (Storer et al., 2021) database to identify repetitive sequences and interspersed repeats were removed from the *bam*files, using *Samtools* . Then, we identified sex-linked scaffolds with the software *SATC* (Nursyifa et al., 2021) and used*Samtools* to remove them. Additionally, we removed reads of mapping quality <30 and regions of low (1/3 mean coverage) and excessive (2x mean coverage) depth, previously estimated with ANGSD (v.10.2.0) (Korneliussen et al., 2014). Since some population genomics analyses can be affected by the presence of first-degree relatives, we calculated relatedness statistics with the software *NgsRelate*(v.2.1) (Hanghøj et al., 2019), which uses genotype likelihoods as input. Subsequently, we removed one sample from the only pair of first-degree relatives found in our dataset from the downstream analyses.

2.3 SNP calling: genotype likelihoods and genotype calls

We called SNPs in two different ways: calculating genotype likelihoods with *ANGSD* and calling genotypes (generation of a *vcf* file) with *Bcftools* . Genotype likelihoods take into account genotype uncertainty and allow to obtain SNPs at very low coverages (Lou et al., 2021). We calculated genotype likelihoods in two datasets, one including Black Sea (BLS) individuals and another without BLS individuals using the *Samtools* model (GL 1), keeping SNPs with a minimum minor allele frequency (MAF) of 0.05, having data in a minimum 75% of the individuals and a SNP p-value<$1e^{-6}$. The beagle file generated with*ANGSD* was used as an input in the population structure analysis and to calculate population genomics summary statistics. Genotypes were called with *Bcftools* commands *mpileup* and *call,*with the multiallelic and rare-variant calling option *-m* , in alignments with minimum mapping *(-q)* and minimum base *(-Q)*quality of 30. We also used *Bcftools* to subsequently retain only high-quality SNPs: we removed non-biallelic sites, indels, SNPs with MAF below 0.05 and SNPs for which we did not yield genotype information in at least 75% of the individuals. The *vcf* file generated was used as an input in the demographic history and seascape genomics analysis.

2.4 Population structure analysis

We studied the genetic structure of North Atlantic harbour porpoises by performing PCAs and admixture analyses on a set of unlinked SNPs. We used the software *ngsLD* (Fox et al., 2019) to prune our data of linked SNPs, considering that SNPs are in linkage disequilibrium (LD) if they are on the same chromosome/scaffold within less than 20 kb and using a minimum weight of 0.5. The PCAs were calculated with*PCAngsd* (Meisner & Albrechtsen, 2018), while the admixture analyses were run in *NGSAdmix* (Skotte et al., 2013). To assess convergence, we performed 20 independent runs, with the number of assumed populations (K) ranging from 2 to 8, a minimum tolerance for convergence of $1x10^{-10}$ and a minimum likelihood ratio value of $1x10^{-6}$.

To remove the genetic signal created by other subspecies and study the population structure of the *P. p. phocoena* subspecies, we excluded BLS samples and repeated the PCA and admixture analysis, with K ranging from 2 to 7. In addition, we examined the local population structure in the Baltic region by including only samples from NOS, BES, and PBS in an independent PCA and admixture analysis, with K ranging from 2 to 3.

2.5 Population genomic summary statistics

A series of diversity and demographic statistics were estimated from the folded site frequency spectrum (SFS) with *ANGSD* . Genome-wide heterozygosity was estimated per sample by first computing the folded site allele frequency likelihood using the reference genome as ancestral state and then calculating the folded SFS. The folded SFS was calculated independently for each sampling site after removing admixed (less than 70% ancestry to any cluster under K4) and migrant individuals (individuals whose ancestry was different from the prevalent cluster of their sampling location). Then we estimated both Watterson's theta and Tajima's D, using a sliding-window approach with window size of 50 kb and a step size of 10 kb. Individual inbreeding coefficients (F) were estimated with the software *ngsF (v1.2.0)* (Vieira et al., 2013). First, approximate F were obtained in an initial run using the *-aprox_ EM* method, with a maximum root mean squared difference between iterations of $1x10^{-5}$ *(-min_ epsilon*) and random initial values. From the output of this first run, the initial parameters for the final run were derived, where the *-min_ epsilon* value was decreased to $1x10^{-7}$ to assume convergence. To avoid convergence to local maxima, this two-step analysis was repeated ten times, as suggested by the authors (Vieira et al., 2013).

2.6 Demographic history analysis

To reconstruct historical relationships among North Atlantic Harbour porpoises, we inferred a maximum likelihood bifurcating population tree using *Treemix* (v1.13) (Pickrell & Pritchard, 2012). The genotypes (*vcf*) were filtered to retain only sites with no missing data. We performed the *Treemix* analysis 1,000 times at the population level with a window size of 1,000 SNPs, specifying the Black Sea subspecies as outgroup and with the option-*noss* to turn off sample size correction, as suggested by the authors. We obtained a consensus tree and bootstrap values with the R package *BITE* (Milanesi et al., 2017).

We also inferred changes in effective population sizes (N_e) through time with the software *SMC++*(Terhorst et al., 2017). We randomly chose three individuals per population, except for the Iberian porpoises, where only one individual was selected for the analysis as the other IBE porpoise was identified as a migrant. This analysis was run only using scaffolds larger than 1Mba (160 scaffolds) and no MAF filtering was applied on the *vcf* . As repetitive and excessive coverage regions were removed from the*bam* files, these uncalled regions were marked as missing data, as suggested by the authors (Terhorst et al., 2017). These regions could be misidentified as very long runs of homozygosity, erroneously decreasing the N_e estimate, and hence compromising the power to infer true population contractions. In addition, we formed composite likelihoods by varying the distinct individual (*-d*) among the three individuals, also as suggested by the authors. Then, the population size histories were computed by using the option*estimate* with the default settings, a generation time of 11.9 years (Taylor et al., 2007) and a mutation rate of $2.56x10^{-8}$ (Yim et al., 2014).

2.7 Genotype-environment associations: Seascape Genomics

To assess how environmental variables shaped the genetic structure of North Atlantic harbour porpoises, we developed a seascape genomics approach. Specifically, we studied genotype-environment associations (GEA) to identify putative SNPs underlying local adaptation by carrying out a redundancy analysis (RDA) with the R package *vegan*(Oksanen, 2011). RDA is a multivariate method that finds linear dependencies between response (genotypes) and explanatory variables (environmental predictors). Previous authors (Forester et al., 2018, Capblancq et al., 2018) have found that multivariate GEA analyses, especially constrained ordination approaches such as RDA, detect more effectively multilocus selection than univariate methods and have a superior combination of low false-positive and high true-positive rates (Grummer et al., 2019; Capblancq & Forester 2021). RDA was carried out at the individual level using two datasets, with and without BLS. To control for spatial autocorrelation and other demographic processes, pairwise oceanic distances were calculated from the coordinates of the samples with the R package*marmap* (Pante & Simon-Bouhet, 2013). Subsequently, pairwise oceanic distances were transformed to distance-based Moran's eigenvector maps (*db-MEMs*) with the R package *adespatial* (Dray et al., 2012) and used as the space variable in the RDA.

Five environmental variables previously identified as significantly associated with genomic variation of Common (*Delphinus delphis*) (Barceló et al., 2022) and Bottlenose dolphins (*Tursiops truncatus*) (Pratt et al., 2022) were selected as potential predictors of harbour porpoise genomic variation: sea surface temperature

(SST), sea surface salinity (SSS), sea current velocity (SCV), sea chlorophyll-A concentration (SCA), and sea primary productivity (SPP). For each of the variables, the annual maximum, mean, minimum and range values were downloaded from the database *BioOracle (v2.2)*(Tyberghein et al., 2012, Assis et al., 2018). SCVmax and SPPmax were not available in this database, yielding altogether 18 variables. As genomic input we used a set of unlinked called genotypes, identified by*Plink (v1.9)* (Purcell et al., 2007) with a window size of 20kb, a step size of 10kb and a r^2 threshold of 0.1.

To maximize the genetic variance explained by our set of environmental predictors we performed a forward selection with the function*forward.sel* of the R package *adespatial* . The forward selection was carried out for the environmental variables and*dbMEMs* separately, and only variables explaining a significant proportion ($p<0.05$) of the genomic variation were retained. To control for multicollinearity, only predictors with a highly conservative variance inflation factor (VIF)<3 were retained. An RDA-based variance partitioning (Capblancq & Forester, 2021) was performed to estimate the independent contribution of the environmental and spatial variables. First, we ran an unconstrained RDA with the retained environmental and spatial variables as predictors to find the variance explained by the full model. Then, we calculated the variance explained by the environmental variables once the influence of spatial variables had been removed by applying a constrained pRDA. Finally, we checked the significance of the RDAs and pRDAs as well as the significance of the environmental variables by an analysis of variance (ANOVA) with 1,000 permutations.

Candidate loci under selection were identified by applying a 3 standard deviation cut-off ($p=0.0027$) on the SNPs loading scores (Forester et al., 2018) from the two first redundancy axis and we determined to which environmental variable each candidate SNPs is most associated. To further control for false positives, we carried out an additional selection genome scan based on individual genotypes with the R package*Pcadapt* (Privé et al., 2020). We used four principal components (K), identified by computing a scree plot and choosing the K that minimized the genomic inflation factor (GIF). Then we transformed the p-values in q-values with the R package *qvalue* and applied a false discovery rate of 0.1 to control for false positives.

To identify candidate genes under local adaptation, we used only the candidate SNPs detected by all three methods RDA-pRDA-Pcadapt on the dataset without BLS and that were associated to salinity (271). We extracted 300 bp of flanking sequences for each candidate SNP resulting in 601 bp long sequences, each containing a single SNP. We then performed a basic local alignment with *BLASTN* using the nucleotide database of NCBI with an e-value of $1x10^{-3}$. After extracting hits per query, we identified gene function and associated biological function with the PubMed and GeneCards (Stelzer et al., 2016) databases.

2.8 Neutral and Adaptive diversity

To get further insights into how adaptive processes shaped the genetic structure of North Atlantic Harbour porpoises, we analyzed separately the neutral and inferred adaptive SNPs. The neutral dataset consisted of the SNPs not found under selection in the *Pcadapt* genome scan and that were in Hardy-Weinberg Equilibrium (HWE), while the adaptive dataset comprised the outlier SNPs identified by the *Pcadapt*genome scan. We performed a PCA with the R package *adegenet*(Jombart & Ahmed, 2011) and estimated mean pairwise-weighted Fst across populations using *vcftools (v0.1.16)* (Danecek et al., 2011). In addition, a mantel test was run with the R package *vegan* to test for isolation by distance (IBD) on both the neutral and adaptive SNP datasets.

3. Results

To disentangle the evolutionary history and population structure of North Atlantic harbour porpoises we resequenced genomes of 74 specimens from 8 different regions (Figure 1). Sample biological information and sequencing statistics are provided in Table S1. After all filtering steps, one BES individual (Figure S1C) kept a small fraction of the raw reads (6.65%, 2.73X coverage) and was discarded from further analysis. Mean sequencing depth was 10.11X (Figure S1B), which is considered an intermediate coverage level (Bourgeouis & Warren, 2021; Fuentes-Pardo & Ruzzante, 2017), and similar depth levels were used in other population genomic studies on cetaceans based on whole genome resequencing (Cerca et al., 2022; de Greef et al., 2022; Zhou et al., 2018). Scaffold coverage comparisons between male and female sequencing data (Figure

S2) identified 47 sex associated scaffolds (122.3 Mb), which were subsequently removed from downstream analysis. Relatedness analysis (Figure S3) identified one pair of first-degree relatives, thus we removed the individual (B41-14) with the lower coverage out of the pair, leading to a final sample size of 72 harbour porpoises. After SNP calling and filtering, we retained a set of 6,186,462 high quality SNPs, from which 1,320,367 were identified as unlinked.

3.1 Genetic structure

The population structure analysis, *PCAngsd* and *NGSAdmix*(Figure 2), based on a set of ˜1.7 million genotype likelihoods, indicated that Harbour porpoise subspecies and populations clustered together. The PCA including all samples (Figure 2A) show two major axes of differentiation: subspecies (PC1) and populations of the North Atlantic subspecies (PC2). The first principal axis, explaining 5.8% of the variance, separates the BLS and one IBE individual from the rest of the samples, suggesting that this IBE individual could belong to the Iberian subspecies. The second principal component (2.2% explained variance) divides Baltic Sea porpoises (BES and PBS) from the Atlantic porpoises (CA, ICE, BAS, NOS), except for one porpoise bycaught in Latvian waters that clustered with the Atlantic ones. Nine samples (1 NOS, 4 BES and 4PBS) were located between the Atlantic and Baltic clusters, but were more closely related to the Baltic one (Figure 2A). When analyzing the PCA without BLS samples (Figure 2C), PC1 (3% variance explained) also divides the Baltic from the Atlantic samples, but in this case only 6 individuals were located between the two groups, as the other three PBS samples were separated by PC2 (2% variance explained). The PCA including only porpoises from the Baltic region (Figure S4A) further suggested the presence of three populations in this small area. The admixture results were consistent with the population structure identified in the PCAs. On the dataset including BLS samples (Figure 2B), K2 separated BLS porpoises from the others, K3 subdivided the Baltic from the Atlantic and K4 isolated the same three PBS samples from the remainder of the Baltic. The admixture analysis without BLS (Figure 2D) and only Baltic region samples (Figure S4B) first divided Baltic samples from the rest (K2), and then the same three PBS individuals from the remaining of the Baltic (K3).

Based on the genetic structure results (Figure 2A,B), and published telemetry and passive acoustic data, we chose K=4 as the highest level of structure and assigned individuals to any of the four clusters if the likelihood of membership was [?] 75% in the admixture analysis. 66 out of the 72 individuals could be assigned to one of the four clusters, leaving six individuals that were admixed between Atlantic and BES cluster. This division resulted in 26 porpoises assigned to the Atlantic population (red cluster), 6 admixed between Atlantic and BES, 32 assigned to the BES population (dark blue cluster), 3 assigned to the Proper Baltic Sea population (light blue cluster) and 5 assigned to the BLS subspecies (black cluster). Both IBE samples clustered with the Atlantic. For the analysis at the population level (*Treemix* ,*SMC++* and population summary statistics), we excluded the 6 specimens admixed between Atlantic and BES, as well as 11 individuals found in the PBS region, but assigned to the Atlantic (red) or BES (dark blue) cluster, such that PBS is represented by the three specimens forming the PBS-specific light blue cluster in Figures 2B and 2D.

3.2 Evolutionary and demographic history

We explored the post-glacial expansion of North Atlantic harbour porpoises by inferring the evolutionary relationships among populations, quantifying their genetic diversity, and estimating historical variation in effective population size (*Ne*). *Treemix* results (Figure 3A) show how the harbour porpoise colonized and dispersed across the North Atlantic from a southern refugium. The first to split were IBE porpoises, followed by NOS. Thereafter, the ancestral porpoise Atlantic subspecies diverged in Baltic populations and the rest of the Atlantic. All internal nodes had a high bootstrap support (90-100%), except the branch with NOS porpoises (71%) which presented a lower support, and the ancestral node of CA and BAS localities (49%) which was statistically not supported, such that the ancestry among the northern North Atlantic porpoises of BAS, CA, and ICE was not resolved in our analysis.

The *SMC++* results (Figure 3B) indicate that overall North Atlantic Harbour porpoises experienced a population expansion and that the N_e of both BLS and IBE subspecies was lower. BLS porpoises' inferred

N_e was stable until ~100,000 years before present (yBP), followed by a steady population contraction up to ~25,000yBP, when theN_e started to increase. IBE porpoises had a very similar trajectory, except that inferred population sizes were slightly higher. The Atlantic and Baltic populations had a related inferred demographic history, with an expansion from ~250,000yBP until ~75,000yBP, when populations trajectories started to diverge. Around 25,000yBP, the Atlantic population (CA, ICE, BAS, NOS) experienced a population contraction, while BES and PBS populations expanded.

Individual genome-wide heterozygosity estimates (Figure 4A) were lower in BLS porpoises compared to the other porpoises. The Iberian individual potentially belonging to a different subspecies also presented a low level of heterozygosity. Among *P.p.phocoena* porpoises, heterozygosity estimates were slightly lower in both populations of the Baltic Sea, compared to the Atlantic sampling locations. Individual inbreeding coefficients (F) were high in BLS porpoises (0.32-0.34) and in the porpoise potentially belonging to*P.p.meridionalis* subspecies (~0.2) (Figure 4A). Atlantic and Baltic populations had inbreeding coefficients close to zero, apart from one individual from PBS that presented a higher$F$ (0.06). Regarding Watterson's theta estimates (Figure S5), the BLS and IBE subspecies as well as the PBS population exhibit low diversity, compared to the BES population and Atlantic locations. Consequently, Tajima's D estimates (Figure 4B) were negative in all regions, except for the PBS population and the BLS subspecies, where they were positive. No Tajima's D was calculated for the Iberian individual, because of too low sample size (n=1).

3.3 Seascape Genomics

Out of the 18 environmental predictors used in the GEAs, forward selection analysis identified five variables significantly associated (p =0.001) with genomic variation (Table S3): mean SST, mean SSS, minimum SCV, minimum SCA, and minimum SPP. The five variables show a heterogeneous seascape in the North Atlantic (Figure S6-S10), especially a pronounced salinity gradient in the Baltic region (Figure S7). Forward selection analysis identified three *dbMEMs* that explained a significant proportion (p <0.05*)* of the genomic variation and were used as the spatial variables. After checking for multicollinearity (Figure S11) and assessing the variance inflation factor (VIF), the five environmental variables were retained since they presented a r<|0.7| and VIF<3. We removed one *dbMEM* in the dataset including BLS, as it had a large VIF (5.8). In total, the RDA model comprised five environmental variables and two *dbMEMs* for the dataset with BLS samples and the five environmental variables and three *dbMEMs* for the dataset excluding BLS (Table S2).

In the dataset without BLS, the overall RDA model was significant (p =0.001), with the environmental variables explaining ~8% of the variation and the spatial variables ~5%. In the RDA model, SSS, SST (p <0.001) and SCV (p =0.023) were significant, while on the pRDA only SSS (p <0.001) and SCV (p =0.018) were significant. By plotting both the RDA and pRDA (Figure 5, Figure S12) we observed that using spatial variables as a condition (pRDA) affected the pattern of the biplots, making the first axis less predominant. In the dataset without BLS, RDA1 (Figure 5A) explained 28% of the variance while pRDA1 (Figure 5B) explained 22.7%. The RDA biplots show the variation in the genomic response to the different environmental variables among sampling locations in the North Atlantic. Both RDA1 and pRDA1 divided the Baltic samples from the rest, mostly based on SSS, while RDA2 and pRDA2 were moderately driven by SST and SCV (Figure 5). In the model including BLS porpoises, the five environmental variables were significant and explained 8.7% of the variance, while the spatial variables explained 4.3% (Table S4). Both pRDA1 and RDA1 (Figure S12) separated BLS from the rest based on SST; pRDA2 and RDA2 (~22% variance explained) separated BES and PBS populations from the rest based on SSS.

The PCA loadings of *Pcadapt* (Figure S13C) showed that most of the p-values followed a uniform distribution, but there was an excess of small p-values, indicating the presence of outliers. Using the dataset without BLS, *Pcadapt* identified 18,955 candidate SNPs, while the pRDA and RDA, identified 9,272 and 7,079 candidate SNPs, respectively (Table 1). A set of 952 candidate SNPs overlapped in the pRDA, RDA and*Pcadapt* . The number of candidate SNPs inferred to be under selection on the dataset with BLS are found in Table S4. We successfully mapped and annotated 202 out of the 271 candidate SNPs associated with salinity, of which 106 were annotated to known genes. While 48 candidate SNPs had hits to only one gene, the other 58 candidate SNPs had equally good (very low e-value and high bit score) hits to multiple annotated genes

(Table S5), thus the latter candidate genes must be interpreted with caution.

3.4 Neutral - adaptive population structure and genetic differentiation

Neutral and adaptive SNP datasets revealed a similar population structure (Figure 6, S14). On the dataset without BLS, both PCAs of the neutral (Figure 6A) and adaptive SNPs identified with *Pcadapt*(Figure 6B) separated the three North Atlantic populations, i.e., Atlantic, BES and PBS. The only difference was that in the neutral set, the three PBS porpoises were separated by PC2 (1.87%) and in the adaptive set by PC1 (15.4%). As to the dataset with BLS, the PCA with the set of neutral SNPs (Figure S14A) separated BLS individuals (PC1) and Baltic from Atlantic porpoises (PC2). The PCA inferred with the set of adaptive SNPs (Figure S14B) was however slightly different: PC1 also separated BLS porpoises from the rest, but PC2 divided the three PBS samples (light blue cluster in Figure 2) from the rest.

Genome-wide pairwise *Fst* levels on the neutral dataset were moderately low (Figure 6C), ranging from ~ 0 to 0.13.*Fst* levels on the adaptive dataset were higher, ranging from ~ 0 to 0.37. Among Atlantic sampling locations (CA, ICE, BAS and NOS), the adaptive dataset presented slightly lower values than the neutral dataset. However, among Baltic (BES and PBS) populations and between Baltic and Atlantic locations, *Fst* was higher in the adaptive dataset than in the neutral dataset. The Mantel tests analyzing associations of pairwise-*Fst* and distance (Figure S14C) among sampling locations were not significant, neither for the neutral nor for the adaptive dataset, as could be expected from the inferred genetic homogeneity across the large geographic range of the open North Atlantic.

4. Discussion

4.1 Population structure and evolutionary history of Harbour porpoises in the North Atlantic and adjacent waters

By using whole-genome resequencing data, as previous authors suggested (Fontaine et al., 2007; 2014; Ben Chehida et al., 2021), we found evidence of three harbour porpoise subspecies and at least three populations within the North Atlantic subspecies (Wiemann et al., 2010; Lah et al., 2016). The three subspecies are *P.p. relicta* in the Black Sea, *P.p. meridionalis* in Iberia and *P.p. phocoena*in the rest of the Atlantic. Black Sea porpoises seem to be a relict population from past glaciations based on the suggested lack of genetic exchange with the other subspecies studied (Figure 2A,B). Our sample set included two samples from Iberian waters, but only one of them was genetically distinct and hence suggestive of belonging to a putative Iberian/Mauritanian subspecies (Fontaine et al., 2007; 2014). PCA results (Figure 2A,C) show that Iberian sample No-2 does not cluster with the Black Sea nor with the Atlantic subspecies. Admixture results grouped this sample with the Atlantic cluster, possibly since*NGSAdmix* does not create a new cluster for only one sample. Nevertheless, while no other sample had any Black Sea ancestry, this sample had a 10% membership to the Black Sea cluster, further suggesting its distinct evolutionary trajectory. More samples from the Iberian/Mauritanian waters should be resequenced to categorically identify these porpoises as an own separate subspecies.

Regarding the population structure within the North Atlantic subspecies,*P.p. phocoena* , we did not find evidence of separation between West and East North Atlantic porpoises, but we confirmed the two proposed distinct porpoise lineages in the peripheral waters of the Baltic Sea (Lah et al., 2016; NAMMCO, 2019). On one hand, PCA, admixture analysis (Figure 2) and Fst levels (Figure 6) show that CA, ICE, BAS and NOS porpoises belong to the same population, the so-called Atlantic population. On the other hand, BES and PBS porpoises clustered separately. Between NOS and BES, there is some gene flow (as indicated by partial assignments to the two respective clusters for some specimens; blue and red in Figure 2). Three very distinct PBS individuals stand out in the genetic structure analysis (Figure 2, 6, S4), which we assigned to the PBS population. This third Baltic cluster was interpreted as the Proper Baltic Sea population for a series of reasons: first, this cluster only comprised porpoises of the PBS region (Figure 2, 6, S14); second, the three porpoises assigned to the PBS population were bycaught during the breeding season, when a separation between putative BES and PBS porpoises occurs (Carlen et al., 2018); third, these three porpoises were bycaught in the easternmost locations (16-18.58W) and in areas known to be important for

the PBS population (Carlen & Evans, 2020), i.e., one in the Gdansk Bay and the other two in the waters surrounding the Swedish island of Oland (Figure S4). These waters present the greatest densities of harbour porpoises during the breeding season in the PBS region (Amundin et al., 2022) and have been reported as potentially important breeding grounds for the PBS population.

The result indicating lack of genetic structure over long distances in the open Atlantic is at odds with the fine-scale population structure we observe in the Baltic region. In an area separated by less than 1,000 kilometers we identified three distinctive lineages (Figure S4C): Atlantic population in the Skagerrak strait, Belt Sea population in the Danish Belts, the Sound and Arkona basin, and the Proper Baltic population in the Baltic proper. We did not find an IBD pattern (Figure S14C), highlighting that geographical distance does not have a major impact on the genomic variation and is unlikely to be a driver of the genetic differences found in North Atlantic harbour porpoises. It should be mentioned, though, that previous authors have found weak patterns of IBD in North Atlantic harbour porpoises (Fontaine et al., 2007; Ben Chehida et al., 2021) when including more sampling locations along the harbour porpoise distribution range. Resequencing more specimens from the open Atlantic may be needed to assess the potential role of geographical distance on North Atlantic harbour porpoise population differentiation.

The demographic history analysis shows that North Atlantic harbour porpoises have been strongly influenced by Pleistocene glaciations, especially since the Last Glacial Period (LGP), when around 25 rapid climate fluctuations occurred until the end of the Last Glacial Maximum (LGM), ~ 19,000yBP (Kindler et al., 2014). TheN_e of the three harbour porpoise subspecies were highly correlated until the onset of the LGP (~110,000yBP) when the Iberian and Black Sea subspecies curves split from the North Atlantic subspecies. Our divergence estimate differs from that of previous authors using a portion of the mitogenome (Fontaine et al., 2014) that dated the most common recent ancestor of North Atlantic subspecies during the LGM. Similar differences in divergence estimates have been reported before in the finless porpoise, when authors using mitochondrial control region data (Wang et al., 2008) inferred a much younger divergence than authors using whole-genome resequencing data (Zhou et al., 2018). The Iberian and Black Sea subspecies originated from the ancestral North Atlantic population when a small group of individuals may have colonized the Mediterranean and Black Sea after the LGP, which left an imprint in the genome as a founder effect. Both Iberian and Black Sea subspecies had lowN_e (Figure 3B), which increased sensitivity to genetic drift which in turn lead to loss of genetic variation, high inbreeding levels (Figure 4A, S5) and a positive Tajima's D (Figure 4C).N_e curves for populations of the North Atlantic subspecies started to diverge around ~75,000yBP, which roughly coincides with two rapid climate fluctuations during that same period (Kindler et al., 2014). This may suggest that BES and PBS lineages have diverged from the Atlantic population during an interglacial period before the formation of the Baltic Sea at the end of the LGM. However, our demographic analysis also indicate that after the LGM, the N_e of the Baltic populations increased. This would be compatible with a scenario in which - as the ice sheets retreated from northern regions at the end of LGM - a group of porpoises colonized the newly formed Baltic Sea, originating the modern BES and PBS populations. Nevertheless, other processes like gene flow and linked selection could be cofounding factors in the SMC++ demographic inferences (Mazet et al., 2016; Schrider et al.,2016), thus these results must be interpreted with caution and time/N_e estimates should not be taken literally.

The maximum likelihood tree inferred with *Treemix* is also compatible with the existence of three subspecies in the North Atlantic and adjacent waters. Previous phylogenies reconstructed with mitogenome data (Ben Chehida et al. 2020) found that BLS porpoises are basal in the North Atlantic, thus we were confident rooting our tree with the BLS subspecies. The first split in our graphical representation of historical relationships among North Atlantic harbour porpoises was the IBE subspecies from the North Atlantic subspecies. Among the North Atlantic subspecies, *Treemix* analysis placed present-day NOS harbour porpoises as basal, compatible with a northward post-glacial expansion from a southern refugia. As the ice sheets retreated, the ancestral North Atlantic subspecies may have started to colonize novel environments in the Baltic, Iceland, Barents Sea, and Canadian waters.

4.2 Local adaptation of Baltic porpoises to low salinity levels

Species and populations inhabiting highly divergent environments are expected to be under different selective pressures, which could cause each local population to evolve traits that provide an advantage under its local environmental conditions (Kawecki & Ebert, 2004). However, the role of ecological specialization on population differentiation and speciation remains poorly understood (Savolainen, Lascoux & Merila, 2013). This is particularly true for cetacean species, with only a few recent studies attempting to address the genetic basis of local adaptation (Barcelo et al., 2022; Pratt et al., 2022; de Greef et al., 2022; Louis et al., 2021; Zhou et al., 2018). Ecologically and geographical marginal environments often host populations at the edge of the species distribution and under extreme selection regimes (Johannesson & Andre, 2006). Examples of such populations are the Belt Sea and Proper Baltic Sea harbour porpoise populations that occur in the peripheral waters of the Baltic Sea, separated from the North Sea by a pronounced salinity gradient.

The GEA results show that including BLS samples in the RDA had a major impact on the outcome of the analysis, especially on the number of putative SNPs inferred to be under selection. The genomic variation of BLS porpoises was highly associated with high temperature (Figure S12), with ~6,000 SNPs correlated with SST and only a few associated with the other variables (Table S4). As water temperatures in the Black Sea are significantly higher than in the rest of the locations (Figure S6) and BLS porpoises are highly divergent, we could not discern whether these ~6,600 SNPs were indeed associated with temperature or were rather very distinct for different evolutionary pressures or pronounced genetic drift in BLS porpoises. Thus, to identify candidate genes associated with environmental variables we focused on the dataset without BLS, where the population divergence was not as strong.

Our seascape genomics analysis provides statistical support for an influence of salinity on population differentiation in the Baltic Sea (Figure 5, S12). SSS was highly significant in both RDA and pRDA, while SST was significant only in the RDA. The salinity gradient could have contributed to the origin of a soft barrier between the Atlantic and the Baltic, leading to adaptive divergence in BES and PBS porpoises, as seen in other cetaceans as the finless porpoise (Zhou et al., 2018). From 272 inferred candidate SNPs, 107 were annotated with genes potentially associated with the salinity gradient in the Baltic (Table S5). We identified three solute carrier group (SLC) genes, a group of membrane transport proteins. Particularly interesting was the SLC10A1 gene, a sodium ion transport with a critical role in the osmoregulation of bile acids in the liver (Kubitz & Hausinger, 2007). We also obtained hits in osmoregulatory genes (AQP9, DYNC2H1) previously inferred to be under selection in cetaceans (Xu et al., 2013, Sao Pedro et al., 2015, Zhou et al., 2018). Further steps to evaluate the candidate SNPs and genes should include the annotation of Gene Ontology (GO) terms as well as genes within 20kb of the candidate SNP (Pratt et al., 2022).

4.3 Implication on species conservation and future perspectives

The ongoing biodiversity crisis is impacting many organisms across the tree of life, and cetaceans are no exception. From the 132 cetacean species, subspecies or populations assessed by the International Union for Conservation of Nature (IUCN), only 51 are considered of least concern, while 24 are considered critically endangered and 24 are classified as endangered. In this study, we have analyzed whole-genome sequencing data of a critically endangered population, the Proper Baltic Sea harbour porpoise, and an endangered subspecies, the Black Sea harbour porpoise. Our results provide genomic evidence that the Proper Baltic Sea population is a distinct lineage (Figure 2, S4) and, given the poor status of the population, urgent measures to protect the species must be implemented (Carlen, Nunny & Simmonds, 2021). Previous studies have shown that microsatellite data does not yield enough statistical power to identify the fine population structure of the harbour porpoise in Baltic waters (Wiemann et al., 2010; Lah et al., 2016). Thus, our results highlight that whole-genome re-sequencing is a powerful tool to unravel even the most subtle population structure. Although in recent years the costs of WGS have been greatly reduced (Fuentes-Pardo & Ruzzante, 2017), it is still prohibitive for small/medium research groups to sequence hundreds or thousands of individuals. In our data, we discovered highly informative SNPs that differentiate the three populations occurring in Baltic waters and which could be used to design a small SNP panel to genotype thousands of samples at a moderate price. Such an approach has been previously implemented in plants (Nygaard et al., 2022), terrestrial (von Thaden et al., 2017) and marine species (Jenkins et al., 2019). With such a panel, bycatch and stranded

harbour porpoises could be genotyped and assigned to populations to monitor the conservation status of Baltic porpoise populations.

We identified high levels of inbreeding in the Iberian subspecies, as well as low genetic diversity and N_e compared with the North Atlantic subspecies. Abundance surveys have estimated the Iberian population to around 2,900 animals and have presented one of the lowest population densities on the European continental shelf (Hammond et al., 2013). Previous studies have reported gene flow from Iberia to more northern regions, but not from the North Atlantic subspecies to Iberian porpoises (Ben Chehida et al., 2021). Therefore, following propositions of previous authors (Fontaine, 2014; Fontaine et al., 2016) we confirm the distinctiveness of Iberian porpoises which may warrant subspecies status. Notwithstanding these taxonomic considerations, measures to guarantee the survival of harbour porpoises in Iberian waters are needed. Similarly, Black Sea porpoises presented high levels of inbreeding, low genetic diversity and N_e , which imply that Black Sea porpoises are subject to demographic stochasticity due to strong genetic drift (Palstra & Ruzzante, 2008). Although there are no reliable estimates of current population size across the entire Black Sea, Harbour porpoise mortality in the Black Sea is high, with thousands of animals each year incidentally bycaught (Birkun and Frantzis, 2008; ACCOBAMS, 2020a). In addition, Harbour porpoises and other cetaceans in the Black Sea are highly affected by activities related to fossil fuels extraction, construction work (as the Kerch Bridge), underwater explosions, and different sources of pollution (Carlen, Nunny & Simmonds, 2021). Thus, explicit management policies must be implemented to protect Black Sea Harbour porpoises.

Acknowledgements

We thank Thorvaldur Gunnlaugsson, Gisli Vikingsson (+), Yves Morin and Santiago Lens for providing access to Icelandic, Canadian and Iberian porpoise samples, respectively. We would further like to thank Katja Havenstein, Ljerka Lah, and all the other people in our lab that isolated DNA from porpoise samples throughout the years. Large-scale computations were carried out at the High-Performance Computing Cluster managed by ZIM (Zentrum fur Informationstechnologie und Medienmanagement) at the University of Potsdam. We are grateful to the ZIM team for assisting with the installation of modules in the computing cluster. This work was financially supported by the Bundesamt fur Naturschutz (FKZ 3521532214) granted to R.T. and the University of Potsdam

Bibliography

Amundin, M., Carlstrom, J., Thomas, L., Carlen, I., Koblitz, J., Teilmann, J., Tougaard, J., Tregenza, N., Wennerberg, D., Loisa, O., ... Benke, H. Estimating the abundance of the critically endangered Baltic Proper harbour porpoise (*Phocoena phocoena*) population using passive acoustic monitoring. (2022). *Ecology and Evolution* , 12:e8554. https://doi.org/10.1002/ece3.8554

Alfonsi, E., Hassani, S., Carpentier, F.-G., Le Clec'h, J.-Y., Dabin, W., Van Canneyt, O., Fontaine, M. C., & Jung, J.-L. (2012). A European melting pot of harbour porpoise in the French Atlantic coasts inferred from mitochondrial and nuclear data. *PLoS One* , *7* (9), e44425. https://doi.org/10.1371/journal.pone.0044425

Assis, J., Tyberghein, L., Bosch, S., Verbruggen, H., Serrao, E.A., de Clerck, O., Tittensor, D. Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modelling. (2018). *Global Ecology and Biogeography* , 27:277-284. https://doi.org/10.1111/geb.12693

Autenrieth, M., Hartmann, S., Lah, L., Roos, A., Dennis, A. B., & Tiedemann, R. (2018). High-quality whole-genome sequence of an abundant Holarctic odontocete, the harbour porpoise (Phocoena phocoena).*Molecular Ecology Resources* , 18(6), 1469–1481. https://doi.org/10.1111/1755-0998.12932

Barcelo, A., Sandoval-Castillo, J., Brauer, C.J., Bilgmann, K., Parra, G.J., Beheregaray, L.B., Moller, L.M. Seascape genomics of common dolphins (Delphinus delphis) reveals adaptive diversity linked to regional and local oceanography. (2022). *BMC Ecology and Evolution* , 22: 88. https://doi.org/10.1186/s12862-022-02038-1

Barret, R.D.H. & Schluter, D. Adaptation from standing genetic variation. (2008). *Trends in Ecology and*

Evolution , 23:38-44. https://doi.org/10.1016/j.tree.2007.09.008

Ben Chehida, Y., Loughnane, R., Thumloup, J., Kaschner, K., Garilao, C., Rosel, P.E., Fontaine, M.C. No leading-edge effect in North Atlantic harbor porpoises: Evolutionary and conservation implications. (2021).*Evolutionary applications,* 14:1588-1611. https://doi.org/10.1111/eva.13227

Ben Chehida, Y., Stelwagen, T.,Hoekendijk1, J.P.A., Ferreira, M., Eira, C., Pereira, A.T., Nicolau, L., Marcalo, A., Thumloup1, J., Fontaine, M.C. Harbor porpoise losing its edges: genetic time series suggests a rapid population decline in Iberian waters over the last 30 years. (2021). *BioRxiv* . https://doi.org/10.1101/2021.08.19.456945

Ben Chehida, Y., Thumloup, J., Schumacher, C., Harkins, T., Aguilar, A., Borrell, A., Ferreira, M., Rojas-Bracho, L., Robertson, K. M., Taylor, B. L., Vikingsson, G. A., Weyna, A., Romiguier, J., Morin, P. A., Fontaine, M. C. (2020). Mitochondrial genomics reveals the evolutionary history of the porpoises (*Phocoenidae*) across the speciation continuum. *Scientific Reports* , 10:15190. https://doi.org/10.1038/s41598-020-71603-9

Berggren, P., Ishaq, R., Zebuhr, Y., Naf, C., Bandh, C., Broman, D. Patterns and Levels of Organochlorines (DDTs, PCBs, non-ortho PCBs and PCDD/Fs) in Male Harbour Porpoises (*Phocoena phocoena*) from the Baltic Sea, the Kattegat-Skagerrak Seas and the West Coast of Norway. (1999). *Marine Pollution Bulletin* , 38:1070-1084.

Birkun Jr., A. 2002. Natural mortality: Black Sea. *Cetaceans of the Mediterranean and Black Seas: State of Knowledge and Conservation Strategies* , pp. 181-193. ACCOBAMS Secretariat, Monaco.

Brownell Jr, R.L., Reeves, R.R., Read, A.J., Smith, B.D., Thomas, P.O., Ralls, K., Amano, M., Berggren P., Chit, A.M., Collins, T., ... Wang, J.Y. Bycatch in gillnet fisheries threatens Critically Endangered small cetaceans and other aquatic megafauna. (2019). *Endangered Species Research* , 40:285-296. https://doi.org/10.3354/esr00994

Bourgeois, Y.X.C. & Warren, B.H. An overview of current population genomics methods for the analysis of whole-genome resequencing data in eukaryotes. (2021). Molecular ecology, 30: 6036-6071. https://doi.org/10.1111/mec.15989

Capblancq, T. & Forester, B.R. Redundancy analysis: A Swiss Army Knife for landscape genomics. (2021). Methods in Ecology and Evolution, 12:2298-2309. https://doi.org/10.1111/2041-210X.13722

Capblancq, T., Luu, K., Blum, M.G.B., Bazin, E. Evaluation of redundancy analysis to identify signatures of local adaptation. (2018). Molecular Ecology Resources, 18:1223-1233. https://doi.org/10.1111/1755-0998.12906

Carlen, I., Nunny, L., Simmonds, M.P. Out of Sight, Out of Mind: How Conservation Is Failing European Porpoises. (2021). *Frontiers in Marine Science* , 8:617478. https://doi.org/10.3389/fmars.2021.617478

Carlen, I. & Evans, P.G.H. Progress report on the Jastarnia plan: the recovery plan for the harbour porpoise in the Baltic Proper. (2020). ASCOBANS/MOP9/Inf.6.1.1b/Rev.2

Carlen, I., Thomas, L., Carlstrom, J., Amundin, M., Teilmann, J., Tregenza, N., ... Acevedo-Gutierrez, A. Basin-scale distribution of harbour porpoises in the Baltic Sea provides basis for effective conservation actions. (2018). *Biological Conservation* , 226:42-53. https://doi.org/10.1016/j.biocon.2018.06.031

Cerca, J., Westbury, M.V., Heide-Jorgensen, M.P., Kovacs, K.M., Lorenzen, E.D., Lydersen, C., Shpak, O.V., Wiig, O., Bachmann, L. High genomic diversity in the endangered East Greenland Svalbard Barents Sea stock of bowhead whales (*Balaena mysticetus*). (2022).*Scientific reports,* 12:6118. https://doi.org/10.1038/s41598-022-09868-5

Chen, S., Zhou, Y., Chen, Y., Gu, J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. (2018). *Bioinformatics* , 34:884.890. https://doi.org/10.1093/bioinformatics/bty560

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Martin O Pollard, M., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., Li, H. Twelve years of SAMtools and BCFtools. (2021). *GigaScience* , 10:1-4. https://doi.org/10.1093/gigascience/giab008

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R., 1000 Genomes Project Analysis Group. The variant call format and VCFtools. *Bioinformatics* , 27: 2156-2158. https://doi.org/10.1093/bioinformatics/btr330

DeFaveri, J., & Merila, J. Local adaptation to salinity in the three-spined stickleback? (2014). *Journal of Evolutionary Biology* , 27:290-302. https://doi.org/10.1111/jeb.12289

de Greef, E., Einfeldt, A.L., Miller, P.J.O., Ferguson, S.H., Garroway, C.J., Lefort, K.J., Paterson, I.G., Bentzen, P., Feyrer, L.J. Genomics reveal population structure, evolutionary history, and signatures of selection in the northern bottlenose whale, *Hyperoodon ampullatus.* (2022). *Molecular Ecology* , 31:4919-4931. https://doi.org/10.1111/mec.16643

Dray, S., Pelissier, R., Couteron, P., Fortin, M.-J, Legendre, P., Peres-Neto, P.R., Bellier, E., Bivand, R., ... Wagner, H.H. (2012). Community ecology in the age of multivariate multiscalespatial analysis. *Ecological Monographs* , 82:257-275. https://doi.org/10.1890/11-1183.1

Dzido, J., Rolbiecki, L., Joanna N. Izdebska a, Jerzy Rokicki a, Tytus Kuczkowski a, Iwona Pawliczka, I. A global checklist of the parasites of the harbor porpoise *Phocoena phocoena,* a critically-endangered species, including new findings from the Baltic Sea. (2021). *International Journal for Parasitology: Parasites and Wildlife* , 15:290-302. https://doi.org/10.1016/j.ijppaw.2021.07.002

Feistel, R., Marion, G.M., , R. Pawlowicz, R., Wright, D.G. Ocean Science Thermophysical property anomalies of Baltic seawater. (2010). *Ocean Science* , 6:949-981. https://doi.org/10.5194/os-6-949-2010

Fontaine, M. C. Harbour Porpoises, *Phocoena phocoena* , in the Mediterranean Sea and Adjacent Regions: Biogeographic Relicts of the Last Glacial Period. (2016). *Advances in marine biology* ,75:333-358. https://doi.org/10.1016/bs.amb.2016.08.006

Fontaine, M. C., Roland, K., Calves, I., Austerlitz, F., Palstra, F. P., Tolley, K. A., Ryan, S., Ferreira, M., Jauniaux, J., Llavona, A., ... Aguilar, A. (2014). Postglacial climate changes and rise of three ecotypes of harbour porpoises, *Phocoena phocoena* , in western Palearctic waters. *Molecular Ecology* , 23(13), 3306-3321. https://doi.org/10.1111/mec.12817

Fontaine, M. C., Baird, S. J., Piry, S., Ray, N., Tolley, K. A., Duke, S., Birkun Jr, A., Ferreira, M., Jauniaux, T., Llavona, A., ... Michaux, J. R. (2007). Rise of oceanographic barriers in continuous populations of a cetacean: the genetic structure of harbour porpoises in Old World waters. *BMC Biology* , 5(1), 30. https://doi.org/10.1186/1741-7007-5-30

Forester, B.R., Lasky, J.R., Wagner, H.H., Urban, D.L. Comparing methods for detecting multilocus adaptation with multivariate genotype-environment associations. (2018). *Molecular Ecology,*27:2215-2233. https://doi.org/10.1111/mec.14584

Formenti, G., Theissinger, K., Fernandes, C., Bista, I., Bombarely, A., Bleidorn, C., Ciofi, C., Crottini, A., ... European Reference Genome Atlas (ERGA) Consortium. The era of reference genomes in conservation genomics. (2022). *Trends in Ecology & Evolution* , 37: 197-202. https://doi.org/10.1016/j.tree.2021.11.008

Foote, A.D., Hooper, R., Alexander, A., Baird, R.W., Baker, C.S., Ballance, L., Barlow, J., Andrew Brownlow, A., Collins, T., Constantine, R., | Luciano Dalla Rosa, L.D., ... Morin, P.M. Runs of homozygosity in killer whale genomes provide a global record of demographic histories. (2021). *Molecular Ecology* , 30:6162-6177. https://doi.org/10.1111/mec.16137

Foote, A.D., Vijay, N., Avila-Arcos, M.C., Baird, R.W., Durban, J.W., Fumagalli, M., Gibbs, R.A., Hanson, B.M., Korneliussen, T.S., Martin, M.D., ... Wolf, J.B.W. Genome-culture coevolution

promotes rapid divergence of killer whale ecotypes. (2016). *Nature Communications* , 7: 11693 https://doi.org/10.1038/ncomms11693

Fox, E.A., Wright, A.E., Fumagalli, M., Vieira, F.G. ngsLD: evaluating linkage disequilibrium using genotype likelihoods. (2019).*Bioinformatics* , 35:3855-3856. https://doi.org/10.1093/bioinformatics/btz200

Fuentes-Pardo, A.P. & Ruzzante, D.E. Whole-genome sequencing approaches for conservation biology: Advantages, limitations, and practical recommendations. (2017). *Molecular Ecology* , 26:5369-5406. https://doi.org/10.1111/mec.14264

Galatius, A., Kinze, C. C., & Teilmann, J. (2012). Population structure of harbour porpoises in the Baltic region: Evidence of separation based on geometric morphometric comparisons. *Journal of the Marine Biological Association of the United Kingdom* , 92:1669-1676. https://doi.org/10.1017/S0025315412000513

Gaskin, D.E. (1992). Status of the harbour porpoise, *Phocoena phocoena* , in Canada. *Canadian Field-Naturalist* , 106: 36-54

Hammond, P. S., Macleod, K., Berggren, P., Borchers, D. L., Burt, L., Canadas, A., ... Vazquez, J. A. Cetacean abundance and distribution in European Atlantic shelf waters to inform conservation and management. (2013). *Biological Conservation* , 164, 107-122. https://doi.org/10.1016/j.biocon.2013.04.010

Hammond, P.S., Bearzi, G., Bjorge, A., Forney, K.A., Karczmarski, L., Kasuya, T., Perrin, W.F., Scott, M.D., Wang, J.Y., Wells, R.S., Wilson, B., 2016. *Phocoena phocoena* Baltic Sea subpopulation (errata version published in 2016) (No. e. T17031A98831650), The IUCN Red List of Threatened Species 2008

Hanghoj, K., Moltke, I., Andersen, P.A., Manica, A., Korneliussen, T.S. Fast and accurate relatedness estimation from high-throughput sequencing data in the presence of inbreeding. . (2019). *GigaScience,* 8:1-9. https://doi.org/10.1093/gigascience/giz034

Hewitt, G. Speciation, hybrid zones and Phylogeography - or seeing genes in space and time. (2001). *Molecular Ecology* , 10: 537-549. https://doi.org/10.1006/bijl.1996.0035

Hewitt, G. The genetic legacy of the Quaternary ice ages. (2000).*Nature* , 405(6789), 907-913. https://doi.org/10.1038/35016000

Hewitt, G. Post-glacial re-colonization of European biota. (1999).*Biological Journal of the Linnean Society* , 68(1-2): 87-112. https://doi.org/10.1006/bijl.1999.0332

Huggenberger, S., Benke, H., & Kinze, C. C. Geographical variation in harbour porpoise (*Phocoena phocoena*) skulls: Support for a separate non-migratory population in the Baltic proper. (2002).*Ophelia* , 56: 1-12. https://doi.org/10.1080/00785236.2002.10409484

Jaramillo-Legorreta, A.M., Cardenas-Hinojosa, G., Nieto-Garcia, E., Rojas-Bracho, L., Thomas, L., Ver Hoef, J.M., Moore, J., Taylor, B., Barlow, J., Tregenza, N. Decline towards extinction of Mexico's vaquita porpoise (*Phocoena sinus*). (2019) Royal Society Open Science, 6(7):190598. https://doi.org/10.1098/rsos.190598

Jenkins, T.L., Ellis, C.D., Triantafyllidis, A., Stevens, J.R. Single nucleotide polymorphisms reveal a genetic cline across the north-east Atlantic and enable powerful population assignment in the European lobster. (2019). *Evolutionary Applications* , 12:1881-1899. https://doi.org/10.1111/eva.12849

Johannesson, K. & Andre, C. Life on the margin: genetic isolation and diversity loss in a peripheral marine ecosystem, the Baltic Sea. (2006).*Molecular Ecology* , 15:2013-2029. https://doi.org/10.1111/j.1365-294X.2006.02919.x

Jombart, T. & Ahmed, I. *adegenet 1.3-1* : new tools for the analysis of genome-wide SNP data. *Bioinformatics* , 27:3070-3071. https://doi.org/10.1093/bioinformatics/btr521

Kardos, M., Armstrong, E.E., Fitzpatrick, S.W., Hauser, S., Hedrick, P.W., Miller, J.M., Tallmon, D.A., Funk, W.C. The crucial role of genome-wide genetic variation in conservation. (2021). PNAS, 118: e2104642118. https://doi.org/10.1073/pnas.2104642118

Karlson, K., Ishaq, R., Becker, G., Berggren, P., Broman, D., Colmsjo, A. PCBs, DDTs and methyl sulphone metabolites in various tissues of harbour porpoises from Swedish waters. (2000). *Environmental Pollution* , , 110:29-46.

Kawecki, T.J. & Ebert, D. Conceptual issues in local adaptation. (2004). *Ecology Letters* , 7:125-1241. https://doi.org/10.1111/j.1461-0248.2004.00684.x

Kindler, P., Guillevic, M., Baumgartner, M., Schwander, J., Landais, A., Leuenberger, M.Temperature reconstruction from 10 to 120 kyr b2k from the NGRIP ice core. (2014). *Climate of the Past* ,10:887-902. https://doi.org/10.5194/cp-10-887-2014

Korneliussen, T.S., Albrechtsen, A., & Nielsen, R. ANGSD: Analysis of Next Generation Sequencing Data. (2014). *BMC Bioinformatics* , 15:1-13. https://doi.org/10.1186/s12859-014-0356-4

Kubitz, R. & Haussinger, D. Osmoregulation of Bile Formation. (2007). Academic Press,Volume 428:313-324. https://doi.org/10.1016/S0076-6879(07)28018-8.

Lah, L., Trense, D., Benke, H., Berggren, P., Gunnlaugsson, Th., Lockyer, C., Ozturk, A., Ozturk, B., Pawliczka, I., Roos, A., Siebert, U., Skora, K., Vikingsson, G., Tiedemann, R., 2016. Spatially Explicit Analysis of Genome-Wide SNPs Detects Subtle Population Structure in a Mobile Marine Mammal, the Harbor Porpoise. PLOS ONE 11, e0162792. https://doi.org/10.1371/journal.pone.0162792

Li, H., & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. (2009). Bioinformatics, 25:1754-1760. https://doi.org/0.1093/bioinformatics/btp324

Lou, R.N., Jacobs, A., Wilder, A.P., Therkildsen, N.O. A beginner's guide to low-coverage whole genome sequencing for population genomics. (2021). *Molecular Ecology* , 30:5966-5993. https://doi.org/10.1111/mec.16077

Louis, M., Galimberti, M., Archer, F., Berrow, S., Brownlow, A., Fallon, R., Nykanen, J., O'Brien, J., Roberston, K.M., Rosel, P.E., ... Gaggioti, O.E. Selection on ancestral genetic variation fuels repeated ecotype formation in bottlenose dolphins. (2021). Science Advances, 7(44): 1245. https://doi.org/10.1126/sciadv.abg1245

Luna D.C.J., Goodman, S.J., Thatcher, O., Jepson, P.D., Andersen, L., Tolley, K., Hoelzel, A.R. (2012). Phenotypic and genetic divergence among harbour porpoise populations associated with habitat regions in the North Sea and adjacent seas. *Journal of Evolutionary Biology,*25: 674-681. https://doi.org/10.1111/j.1420-9101.2012.02461.x

Mazet, O., Rodriguez, W., Grusea, S., Boitard, S., Chikhi, L. On the importance of being

structured: Instantaneous coalescence rates and human evolution-Lessons for ancestral

population size inference? (2016) *Heredity* , 116:362–371. https://doi.org/10.1038/hdy.2015.104

Mei, Z., Zhang, X., Huan, S., Zhao, X., Hao, Y., Zhang, L., Qian, Z., Zheng, J., Wang, K., Wang, D. The Yangtze finless porpoise: on an accelerating path to extinction? (2014). *Biological Conservation* , 172: 117-123. http://dx.doi.org/10.1016/j.biocon.2014.02.033

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

Milanesi, M., Capomaccio, S., Vajana, E., Bomba, L., Garcia, J. F., Ajmone-Marsan, P., Colli, L. BITE: An R package for biodiversity analyses. (2017). *bioRxiv* , 181610 https://doi.org/10.1101/181610

NAMMCO, N. A. M. M. C. and the N. I. of M. R. Report of the Joint IMR/NAMMCO International Workshop on the Status of Harbour Porpoises in the North Atlantic. (2019).

Nursyifa, C., Bruniche-Olsen, A., Garcia-Erill, G., Heller, R., Albrechtsen, A. Joint identification of sex and sex-linked scaffolds in non-model organisms using low depth sequencing data. (2022).*Molecular Ecology Resources,* 22:458-467. https://doi.org/10.1111/1755-0998.13491

Nygaard, M., Kopatz, A., Speed, J.M.D., Martin, M.D., Presto, T., Kleven, O., Bendiksby, M. Spatiotemporal monitoring of the rare northern dragonhead (Dracocephalum ruyschiana, Lamiaceae) - SNP genotyping and environmental niche modeling herbarium specimens. (2022). Ecology and Evolution, 12:e9187. https://doi.org/10.1002/ece3.9187

Oksanen, J. Vegan: community ecology package. R package version 1.8-5. (2007). http://www.cran.r-project.org

Olsen, M.T., Nielsen, N.H., Biard, V., Teilmann, J., Ngo, M.C., Vikingsson, G., Gunnlaugsson, T., Stenson, G., Lawson., Lah, L., Tiedemann, R., PR Heide-Jorgensen, P.R. Genetic and behavioral data confirm the existence of a distinct harbour porpoise ecotype in West Greenland. (2022). *Ecological Genetics and Genomics,* 22:100-108. https://doi.org/10.1016/j.egg.2021.100108

Palstra, F.P. & Ruzzante D.E. Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? (2008).*Molecular Ecology* , 17:3428-3447. https://doi.org/10.1111/j.1365-294X.2008.03842.x

Pante, E. & Simon-Bouhet, B. marmap: A Package for Importing, Plotting and Analyzing Bathymetric and Topographic Data in R. (2013). *PLoS ONE* , 8:e73051. https://doi.org/10.1371/journal.pone.0073051

Pereyra, R.T., Bergstrom, L., Kautsky, L., Johannesson, K. Rapid speciation in a newly opened postglacial marine environment, the Baltic Sea. *BMC Evolutionary Biology* , 9:70. https://doi.org/10.1186/1471-2148-9-70

Pickrell, J.K. & Pritchard, J.K. Inference of population splits and mixtures from genome-wide allele frequency data. (2012). *PLoS Genetics* , 8: e1002967. https://doi.org/10.1371/journal.pgen.1002967

Pratt, E.A.L., Beheregaray, L.B., Bilgmann, K., Zanardo, N., Diaz-Aguirre, F., Brauer, C., Sandoval-Castillo, J., Moller, L.M. Seascape genomics of coastal bottlenose dolphins along strong gradients of temperature and salinity. (2022). *Molecular Ecology* , 31:2223-2241. https://doi.org/10.1111/mec.16389

Prive, F., Luu, K., Vilhjalmsson, B.J., Blum, M.G.B. Performing Highly Efficient Genome Scans for Local Adaptation with R Package pcadapt Version 4. (2020). *Molecular Biology and Evolution* , 37:2153–2154. https://doi.org/10.1093/molbev/msaa053

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., Sham, P.C. PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics,*81:559-75. https://doi.org/10.1086/519795

Quintela, M., Besnier, F., Seliussen, B., Glover, K.A., Lindstrom, U. Population structure of bycaught harbour porpoise *(Phocoena phocoena)* in Norway. (2020). *Marine Biology Research* , 16(3): 141-147. http://dx.doi.org/10.1080/17451000.2020.1729992

Read, A. J. (1999). Harbour porpoise *Phocoena phocoena* (Linnaeus, 1758). In R. J. Harrison, & S. H. Ridgway (Eds.), *Handbook of marine mammals. Second book of dolphins* (Vol. *6* , pp. 323–350). Academic Press

Reckendorf, A., Everaarts, E., Bunskoek, P., Haulena, M., Springer, A., Lehnert, K., Lakemeyer, J., Siebert, U., Strube, C. Lungworm infections in harbour porpoises (Phocoena phocoena) in the German Wadden Sea between 2006 and 2018, and serodiagnostic tests. (2021).*International Journal for Parasitology: Parasites and Wildlife* , 14:53-61. https://doi.org/10.1016/j.ijppaw.2021.01.001

Riginos, C. & Cunningham, C.W. Local adaptation and species segregation in two mussel (*Mytilus edulis* x*Mytilus trossulus*) hybrid zones. (2005) *Molecular Ecology* , 14:381-400. https://doi.org/0.1111/j.1365-294X.2004.02379.x

Robinson, J.A., Kyriazis, C.C., Nigenda-Morales, S.F., Beichman, A.C., Rojas-Bracho, L., Robertson, K.M., Fontaine, M.C., Wayne, R.K., Lohmueller, K.E., Taylor, B.L., Morin, P.M. The critically endangered vaquita is not doomed to extinction by inbreeding depression. (2022).*Science* , 376: 635-639. https://doi.org/10.1126/science.abm1742

Ryeng, K.A., Lakemeyer, J., Roller, M., Wohlsein, P., Siebert, U. Pathological fndings in bycaught harbour porpoises (Phocoena phocoena) from the coast of Northern Norway. (2022). *Polar Biology* , 45:45-57. https://doi.org/10.1007/s00300-021-02970-w

Sao Pedro, S.L., Alves, J.M.P., Barreto, A.S., Lima, A.O.S. Evidence of Positive Selection of Aquaporins Genes from *Pontoporia blainvillei* during the Evolutionary Process of Cetaceans. (2015).*PLoS ONE* , 10:e0134516. https://doi.org/10.1371/journal.pone.0134516

Savolainen, O., Lascoux, M. & Merila, J. Ecological genomics of local adaptation. (2013). *Nature Review Genetics* ,14:807-820. https://doi.org/10.1038/nrg3522

Schrider, D.R., Shanku, A.G., Kern, A.D. Effects of linked selective sweeps on demographic

inference and model selection. (2016). *Genetics* , 204:1207-1223. https://doi.org/10.1534/genetics.116.190223

Schweizer, R.M., Saarman, N., Ramstad, K.M., Forester, B.R., Kelley, J.L., Hand, B.K., Malison, R.M., Ackiss, A.S., Watsa, M., Nelson, T.C., Beja-Pereira, A., Waples, R.S., Funk, W.C., Luikart, G. Big Data in Conservation Genomics: Boosting Skills, Hedging Bets, and Staying Current in the Field. (2021). *Journal of Heredity* , 112:313-327. https://doi.org/10.1093/jhered/esab019

Siebert, U., Sturznickel, J., Schaffeld, T., Oheim, R., Rolvien, T., Prenger-Berninghoff, E., Wohlsein, P., Lakeyer, J., Rohner, S., Schick, L.A., Gross, S. ... Morell, M. Blast injury on harbour porpoises (*Phocoena phocoena*) from the Baltic Sea after explosions of deposits of World War II ammunition. (2022). *Environment International,* 159: 107014. https://doi.org/10.1016/j.envint.2021.107014

Sjoqvist, C., Godhe, A., Jonsson, P.R., Sundqvist, L., Kremp, A. Local adaptation and oceanographic connectivity patterns explain genetic differentiation of a marine diatom across the North Sea-Baltic Sea salinity gradient. (2015). *Molecular Ecology* , 24:2871-2885. https://doi.org/10.1111/mec.13208

Skotte, L., Korneliussen, T.S. Albrechtsen, Anders (2013): Estimating individual admixture proportions from next generation sequencing data.*Genetics* , 195:693-702. https://doi.org/10.1534/genetics.113.154138

Slatkin, M. Gene flow and the geographic structure of natural populations. (1987). *Science* , 236:787-792

Smit, A.F.A., Hubley, R. & Green, P. *RepeatMasker Open-4.0* . (2013).

Stange, M., Barrett, R.D.H. & Hendry, A.P. The importance of genomic variation for biodiversity, ecosystems and people. (2021). *Nature Reviews Genetics* , 22:89-105. https://doi.org/10.1038/s41576-020-00288-7

Stelzer, G., Rosen, R., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., Iny, S.T., Nudel, R., ... Lancet, D. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analysis. (2016). *Current Protocols in Bioinformatics*, 54:1.30.1-1.30.33. https://doi.org/10.1002/cpbi.5.

Storer, J., Hubley, R., Rosen, J., Wheeler, T.J., Smit, A.F.A. The Dfam community resource of transposable element families, sequence models, and genome annotations. (2021). *Mobile DNA* , 12:2. https://doi.org/10.1186/s13100-020-00230-y

Stuckas, H., Stoof, H., Quesada, H., Tiedemann, R. Evolutionary implications of discordant clines across the Baltic Mytilus hybrid zone (*Mytilus edulis* and *Mytilus trossulus*). (2009).*Heredity* , 103:146-156

Sveegaard, S., Galatius, A., Dietz, R., Kyhn, L., Koblitz, J.C., Amundin, M., Nabe-Nielsen, J., Sinding, M.-H.S., Andersen, L.W., Teilmann, J., 2015. Defining management units for cetaceans by combining genetics, morphology, acoustics and satellite tracking. *Global Ecology Conservation* , 3: 839-850

Szarmach, S.J., Brelsford, A., Witt, C.C., Toews, D.P.L. Comparing divergence landscapes from reduced-representation and whole genome resequencing in the yellow-rumped warbler *(Setophaga coronata)*species complex. (2021). *Molecular Ecology* , 30:594-6005. https://doi.org/10.1111/mec.15940

Tanaka Y. Extinction of populations by inbreeding depression under stochastic environments. (2000). *Population Ecology* , 42: 55-62. https://doi.org/10.1007/s101440050009

Taylor, B.L., Chivers, S.J., Larese, J., Perrin, W.F. Generation length and percent mature

estimates for IUCN assessments of cetaceans. (2007). Southwest Fisheries Science Center, Administrative Report LJ-07-01

Terhorst, J., Kamm, J.A., Song, Y.S. Robust and scalable inference of population history from hundreds of unphased whole genomes. (2017).*Nature genetics* ,49:303-309. https://doi.org/10.1038/ng.3748

Tiedemann, R., Harder, J., Gmeiner, C., & Haase, E. (1996). Mitochondrial DNA sequence patterns of harbour porpoises (*Phocoena phocoena*) from the North and Baltic Sea. *Zeitschrift Fur Saugetierkunde* , 61: 104-111

Tolley, K.A., & Rosel, P.E. (2006). Population structure and historical demography of eastern North Atlantic harbour porpoises inferred through mtDNA sequences. *Marine Ecology Progress Series* , 327: 297-308. https://doi.org/10.3354/meps3 27297

Westbury, M.V., Petersen, B., Garde, E., Heide-Jorgensen, M.P., Lorenzen, E.D. (2019). Narwhal genome reveals long term low genetic diversity despite current large abundance size. *iScience,* 15: 592-599. https://doi.org/10.1016/j.isci.2019.03.023

Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F., de Clerck, O. Bio-ORACLE: a global environmental dataset for marine species distribution modelling. (2012). *Global Ecology and Biogeography* , 21:272-281. https://doi.org/10.1111/j.1466-8238.2011.00656.x

van der Auwera, G.A., Mauricio, O.M., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., ... DePristo, M.A. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. (2013). *Current protocols in bioinformatics* 43, 11.10.1-11.10.33. https://doi.org/10.1002/0471250953.bi1110s43

Vieira, F.G., Fumagalli, M., Albrechtsen, A., Nielsen, R. Estimating inbreeding coefficients from NGS data: Impact on genotype calling and allele frequency estimation. (2013). *Genome Research* , 23:1852-1861. https://doi.org/10.1101/gr.157388.113

von Thaden, Alina; Cocchiararo, B., Jarausch, A., Jungling, H., Karamanlidis, A.A., Tiesmeyer, A., Nowak, C., Munoz-Fuentes, V. Assessing SNP genotyping of noninvasively collected wildlife samples using microfluidic arrays. (2017). *Scientific reports* ,7:10768. https://doi.org/10.1038/s41598-017-10647-w

Wang, J., Frasier, T., Yang, S., White, B. Detecting recent speciation events:

the case of the finless porpoise (genus *Neophocaena*). (2008).*Heredity* , 101:145-155. https://doi.org/10.1038/hdy.2008.40

Wennerstrom, L., Laikre, L., Ryman, N., Utter, F.M., Ghani, N.I.A., Andre, C., DeFaveri, J., Johansson, D., Kautsky, L., Merila, J., ... Primmer, C.R. Genetic biodiversity in the Baltic Sea: species-specific patterns challenge management. (2013).*Biodiversity conservation* , 22:3045-3065. https://doi.org/10.1007/s10531-013-0570-9

Wiemann, A., Andersen, L.W., Berggren, P., Siebert, U., Benke, H., Teilmann, J., Lockyer, C., Pawliczka, I., Skora, K., ... Tiedemann, R. Mitochondrial Control Region and microsatellite analyses on harbour porpoise (*Phocoena phocoena*) unravel population differentiation in the Baltic Sea and adjacent waters. (2010). *Conservation Genetics* , 11(1): 195-211. https://doi.org/10.1007/s10592-009-0023-x

Wolf, M., de Jong, M., Halldorsson, S.H., Arnason , U., Janke, A. Genomic Impact of Whaling in North Atlantic Fin Whales. (2022),*Molecular Biology and Evolution* , 39:msac094. https://doi.org/10.1093/molbev/msac094

Wrange, A.L., Andre, C., Lundh, T., Lind, U., Blomber, A., Jonsson, P.J., Havenhand, J.N. Importance of plasticity and local adaptation for coping with changing salinity in coastal areas: a test case with barnacles in the Baltic Sea. (2014). *BMC Ecology and Evolution* , 14:156. https://doi.org/10.1186/1471-2148-14-156

Xu, S., Yang, Y., Zhou, X., Zhou, K., Yang, G. Adaptive evolution of the osmoregulation-related genes in cetaceans during secondary aquatic adaptation. (2013). *BMC Evolutionary Biology* ,13:189. https://doi.org/10.1186/1471-2148-13-189

Yim, H.S., Cho, Y.S., Guang, X., Kang, S.G., Jeong, J.Y., Cha, S.S., Oh, H.M., Lee, J.H.,

E. C. Yang, Kwon, K.K., Kim, Y.J., ... Lee, J.H. Minke whale genome and aquatic adaptation in cetaceans. (2014). *Nature Genetics,*46:88-92. https://doi.org/10.1038/ng.2835

Zhao, X., Barlow, J., Taylor, B.L., Pitman, R.L., Wang, K., Wei, Z., Stewart, B.S., Turvey, S.T., Akamatsu, T., Reeves, R.R., Wang, D. Abundance and conservation status of the Yangtze finless porpoise in the Yangtze River, China. (2008). *Biological Conservation* , 141: 3006-3018. https://doi.org/10.1016/j.biocon.2008.09.005

Zhou, X., Guang, X., Sun, D., Xu, S., Li, M., Seim. I., Jie, W., Yang, L., Zhu, Q., Xu, J., ... Yang, G. Population genomics of finless porpoises reveal an incipient cetacean species adapted to freshwater. (2018). *Nature Communications* , 9: 1276. https://doi.org/10.1038/s41467-018-03722-x

Data availability

Raw sequencing data are available at the NCBI database under accession numbers XXX.

Author contributions

E.C. and R.T. designed the study; M.A. assisted with bioinformatic analysis; R.T. provided funding; A.R., I.P., M.Q., U.L., H.B., U.S., C.L., P.B., A.O., B.O., and V.L. provided samples and associated biological information; E.C. performed laboratory work, executed bioinformatic analyses and analyzed the data; E.C. and R.T. interpreted the results, E.C. wrote the draft manuscript with input from R.T and M.A. All authors edited and approved the final manuscript.

Table 1 : Analysis of variance (ANOVA) assessing the amount of genomic variation explained by the RDA and pRDA models (BLS excluded). Number of candidate SNPs identified with the RDA, pRDA, *Pcadapt*genome scan, and overlapping between RDA, pRDA, and the 18,955 outlier SNPs identified by Pcadapt. Percentage explained of each environmental variable and its significance is also shown. *** highly significant (p <0.001); * significant (p <0.05)

Variable	RDA	%Variance explained RDA	pRDA	% Variance explained pRDA	Overlap
SST	1,083	19.4 ***	1,351	19.4	154
SSS	1,546	26 ***	1,884	22.3 ***	271
SCV	1,092	18.7 *	1,586	20.1 *	141
SCA	2,173	18	2,617	19.1	222
SPP	1,185	17.9	1,834	19.1	164
Total	7,079	100	9,272	100	952

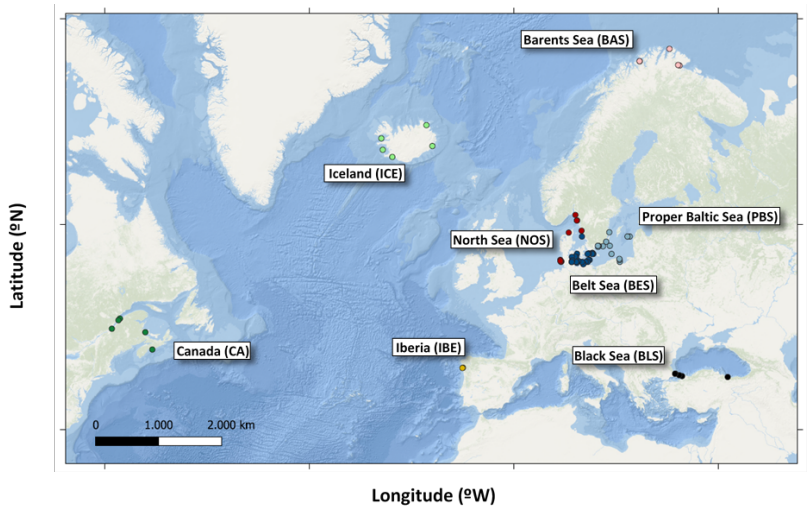

Figure 1 : Map of sampling locations of Harbour porpoise individuals coloured according to origin: Canada (CA), Iceland (ICE), Barents Sea (BAS), North Sea (NOS), Belt Sea (BES), Proper Baltic Sea (PBS), Iberia (IBE), and Black Sea (BLS).

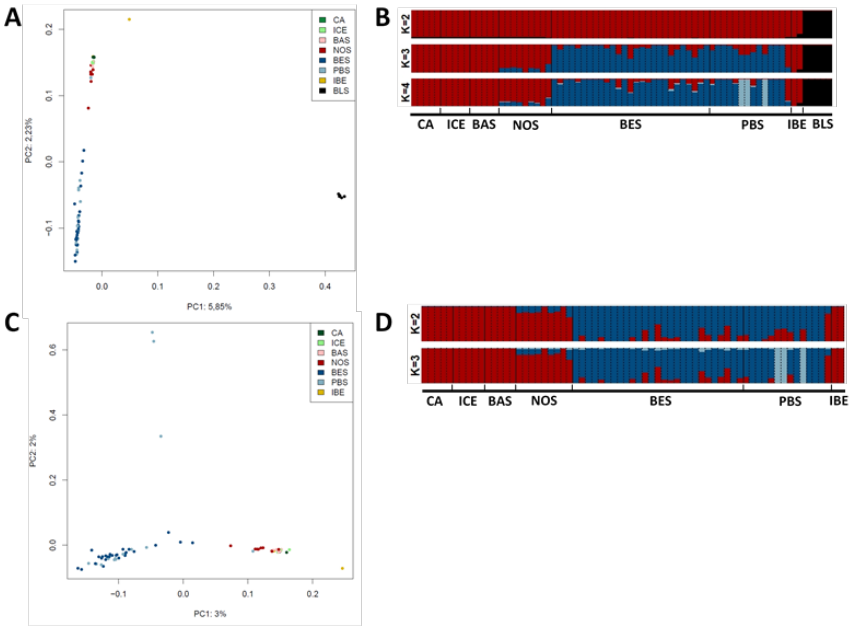

Figure 2 : Population structure of North Atlantic Harbour porpoises indicating the existence of five major genetic clusters: Black Sea subspecies, Iberian subspecies, Atlantic, Belt Sea and Proper Baltic Sea populations. (a) Principal Component Analysis (PCA) of harbour porpoises (N = 72) showing the first and second PCs. (b) Admixture analysis of harbour porpoises (N = 72), only K ranging from 2 to 4 is shown. (c) PCA of the dataset without the Black Sea subspecies (N = 67) showing the first and second PCs. (d) Admixture analysis of the dataset without the Black Sea subspecies (N = 67), only K ranging from 2 to 3

is shown. Each small vertical bar in the admixture analyses represent a Harbour porpoise specimen and the colouring corresponds to its genetic ancestry, Black Sea subspecies in black, Atlantic population in red, Belt Sea population in dark blue and Proper Baltic Sea population in light blue.

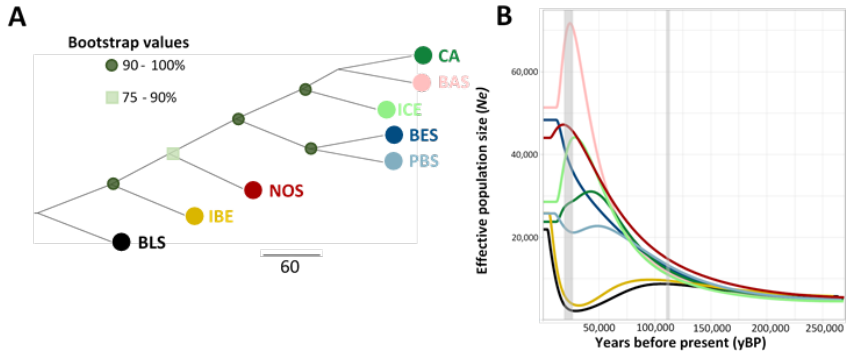

Figure 3 : Historical relationships and demographic history of North Atlantic Harbour porpoises. (a) Maximum likelihood bifurcating tree inferred by *Treemix* indicating a post-glacial colonization of harbour porpoises from a southern refugium. The horizontal branch lengths are proportional to the amount of genetic drift. (b) Inferred changes on effective population size (Ne) through time with a mutation rate of 2.56 x 10^{-8} and a generation time of 11.9 years. The start of the last glacial period (110,000 yBP) and the Last Glacial Maximum (26,500 to 19,000 yBP) are indicated in gray.

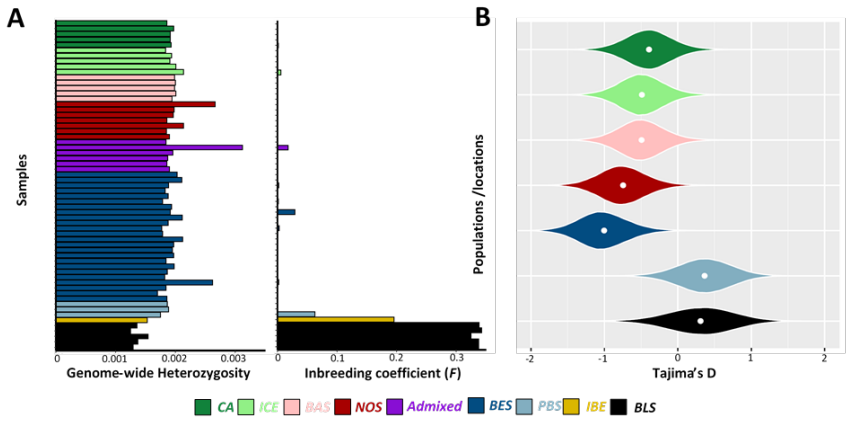

Figure 4 : Population genomics summary statistics of North Atlantic Harbour porpoises. (a) Individual genome-wide heterozygosity, each sample is coloured according to its sampling locality, admixed individuals are coloured in purple and migrants were not included. (b) Individual inbreeding coefficient for each sample. (c) Violin plots of Tajima's D values estimated at population/location level with a sliding-window approach with window size of 50 kb and a step size of 10 kb; the white dot indicates the mean value across the 50 kb Windows. No Tajima's D was calculated for the Iberian subspecies due to small sample size (n =1).

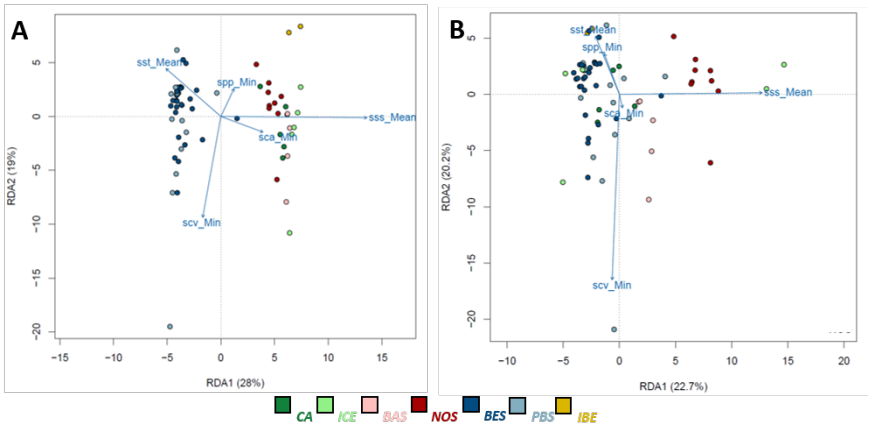

Figure 5 : Genotype-environment association analysis between the five retained environmental variables and a set of 1,320,367 unlinked SNPs in North Atlantic Harbour porpoises, coloured by sampling locality. (a) RDA biplot of the dataset without BLS samples; the overall model was significant (p =0.001), the environmental variables explained 8.5% of the variance. SST, SSS and SCV were significant (b) pRDA biplot of the dataset without BLS, the overall model was also significant (p =0.001), the spatial and environmental variables explained ~5% and ~8% of the variance, respectively. SSS and SCV were significant. sst_Mean (mean sea surface temperature), sss_Mean (mean sea surface salinity), scv_Min (minimum sea current velocity), sca_Min (minimum sea chlorophyll-A concentration) and spp_Min (minimum sea primary productivity).

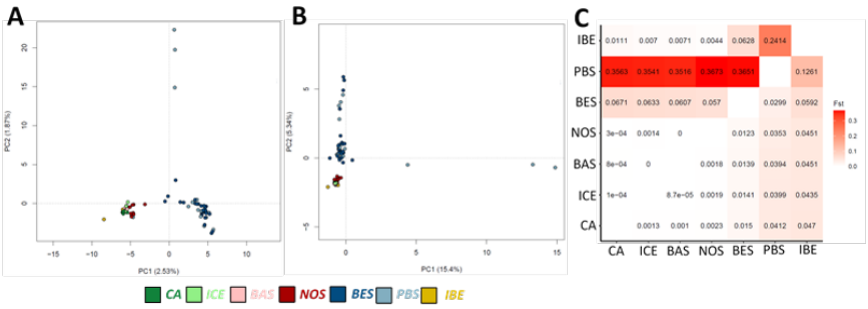

Figure 6 : Genetic structure of neutral and inferred adaptive SNPs on the dataset without BLS. (a) Neutral dataset (1,231,060 SNPs). (b) Adaptive dataset (18,955 SNPs). The three PBS specimens separated out in A and B are those assigned to be PBS population by the*NGSAdmix* and *PCAngsd* results. (c) Heatmap of mean pairwise-weighted Fst across sampling locations calculated with*vcftools* . Fst estimates using the adaptive set of SNPS (18,955) are on the upper left half of the matrix, while the estimates based on the neutral set (1,231,060) are on the lower right half below the diagonal.