ICES Journal of Marine Science, 2025, Vol. 82, Issue 12, fsaf231
https://doi.org/10.1093/icesjms/fsaf231
Received: 4 April 2025; **revised:** 26 November 2025; **accepted:** 1 December 2025
Original Article

Genomic patterns and levels of genetic diversity reveal recent declines in the effective size of European hake populations

Natalia Petit-Marty [1,*], Laura Casas [1], Sara Rocha [2], María Saura [1], Inés González-Castellano [2,3], Sergio Ramirez-Amaro [4,5], Erik Eschbach [6], Beatriz Guijarro [4], Reinhold Hanel [6], Armando Caballero [2,*], Fran Saborido-Rey [1,*]

[1]Institute of Marine Research, CSIC, Eduardo Cabello N° 6, 36208. Vigo, Spain
[2]Centro de Investigación Mariña, Universidade de Vigo, 36310 Vigo, Spain
[3]Universidade da Coruña, Rúa da Maestranza 9, 15001 A Coruña, Spain
[4]Centre Oceanogràfic de les Balears (COB-IEO), CSIC, Moll de Ponent s/n, 07015 Palma, Spain
[5]Laboratori de Genètica, Universitat de les Illes Balears, Ctra. Valldemossa km 7.5, 07122 Palma, Spain
[6]Thünen-Institute of Fisheries Ecology, Herwigstraße 31, 27572 Bremerhaven, Germany

*Corresponding authors. Natalia Petit-Marty, Institute of Marine Research, CSIC, Eduardo Cabello N° 6, 36208. Vigo, Spain. E-mail: npetit@iim.csic.es: Armando Caballero, Centro de Investigación Mariña, Universidade de Vigo, 36310 Vigo, Spain. E-mail: armando@uvigo.es: Fran Saborido-Rey, Institute of Marine Research, CSIC, Eduardo Cabello N° 6, 36208. Vigo, Spain. E-mail: fran@iim.csic.es

Abstract

European hake is an economically important fishery resource with known declines in population abundances and biomass in the last decades. However, little is known about the effect of these abundance declines on the genetic diversity levels of the stocks. In this study, we analysed genomic data of Atlantic and Mediterranean European hake populations, representing the main fishery stocks. The results indicate that abundance declines have likely affected the levels and patterns of genetic diversity within and among the three analysed populations. The Mediterranean population shows lower genetic diversity than the Atlantic ones and can be slightly genetically differentiated from them. Results indicate significant deviations from neutral expectations for constant population sizes, and substantial population declines in all three populations. Estimates of recent and ancestral effective population size (N_e) obtained from coalescence, linkage disequilibrium and sibship methods were coherent with each other. Historical N_e estimates suggest drastic declines of more than 90% in all three populations within the last 15 generations (i.e. nearly 50 years). Positive correlations observed between historical N_e and stock parameters indicative of abundances support the historical trends of estimated N_e. Since genetic variation is essential for populations facing anthropogenically induced environmental changes, this study demonstrates the importance of population genetic estimates for planning management strategies targeting the long-term sustainability of European hake fisheries.

Keywords: hake; Genomics; adaptive potential; climate change; fisheries; genetic diversity; sustainability; conservation

Introduction

Fluctuations in the abundance of wild fish populations are expected to occur due to environmental stochasticity. However, whether and how this natural variation synergises with fishing activities to drive population evolution has yet not been exhaustively investigated (Allendorf et al. 2008, Kuparinen et al. 2016). Population dynamics are determined by the interaction of different factors such as the carrying capacity of the environment, reproductive rates, life-history traits, strategies of population growth, and harvesting rate. Thus, fisheries management needs to consider as much relevant biological and ecological information as possible to ensure long-term sustainability in the use of marine resources (Ovenden et al. 2016). However, knowledge of the relevant biological and ecological factors influencing population dynamics of marine species is fragmentary and methodologically complex to collect. Whereas more than 30% of the fisheries worldwide are exploited at levels considered biologically unsustainable (FAO 2022), anthropogenically induced environmental changes are accelerating (Ceballos et al. 2017). Thus, fish populations are experiencing multiple environmental selective pressures, beyond those of fishing. Hence, it becomes essential to estimate surrogates' indexes such as those offered by population genetics (Ovenden et al. 2016) to forecast population tendencies under the current global scenario of environmental change.

The resilience and potential of adaptation to future environmental challenges of populations largely depend on their amount of genetic diversity (Lynch & Conery 2003, Spielman et al. 2004). This latter is shaped by different factors, such as natural and artificial selection, mutation, migration and genetic drift (Caballero 2020). Fish species subjected to persistent high fishing mortality are expected to undergo population size declines, but also truncated demographies and induced adaptive changes given the selective nature of fishery (Allendorf et al. 2008, Hidalgo et al. 2014). These changes may increase genetic drift, the random change in allele frequencies in finite populations, leading to loss of genetic variation and consequently reduced adaptation capacity and increased vulnerability (Allendorf et al. 2008, Pinsky & Palumbi 2014, Kuparinen et al. 2016, Petit-Marty et al. 2022). A

parameter which quantifies the amount of genetic drift in a population is the effective population size (N_e), the size of an ideal population which shows the same amount of genetic drift and inbreeding as the population under study (Waples 2024a, Wright 1931, Caballero 1994, Wang et al. 2016). Under a neutral model of molecular evolution, the genetic diversity of a population is a function of the mutation rate and N_e (Nei 1987). Thus, a decline in N_e directly implies a loss in genetic diversity and the adaptive potential of the population (Frankham 1995, Spielman et al. 2004, Petit-Marty et al. 2022).

The effective size of a population is strongly related with the reproductive strategy and success, body size, geographic distribution, and generation time length (Lynch & Conery 2003, Wang et al. 2016). It is expected that species with shorter generation times, lower body sizes, and wider distributions tend to have larger N_e than species with narrower distributions, and large body sizes and/or generation times (Lynch and Conery 2003, Manel et al. 2020, Petit-Marty et al. 2022). All these factors directly related to N_e are also relevant to fisheries management strategies. In fishery management, spawning stock biomass (SSB) is often used as an index of the stock reproductive potential (SRP), and hence to predict future recruitment through stock-recruitment relationships, a cornerstone in fisheries management to determine biological reference points informing management decisions. However, it has been largely demonstrated that SSB is not proportional to SRP due to the existence of parental effects (Saborido-Rey and Junquera 1999, Hixon et al. 2014). Additionally, the reproductive success is often attributed to a small portion of the adult population consistent with the concept of sweepstake reproductive success (Hauser & Carvalho 2008, Árnason et al. 2023). In this context, N_e arises as a parameter that provides valuable information on stock sustanaibility (Ovenden et al. 2016), as changes in N_e associated with fishing are mainly due to reduced capacity of the parental stock to produce a sufficient level of recruitment to maintain levels of genetic diversity, whereas changes in abundance depend on both recruitment and mortality (Kuparinen et al. 2016). Thus, fishing may have a large impact on reducing population sizes, productivity and recruitment and, consequently, accelerating the loss of genetic diversity and eroding adaptive potential (Spielman et al. 2004, Pinsky and Palumbi 2014). However, the recovery of abundance is not necessarily accompanied by a parallel recovery of genetic diversity, as the latter is expected to take much longer (Lynch and Conery 2003, Petit-Marty et al. 2022). This is particularly relevant for species with large effective population sizes (N_e), such as for marine species (Clarke et al. 2024), because the time required to recover genetic diversity after a bottleneck is proportional to N_e.

In practice, as N_e is strongly affected by declines in population abundance, an assessment of whether recovery of depleted stocks has restored the population's genetic diversity levels can be made by comparing N_e estimates before and after abundance depletion. Historical changes in N_e for long periods of time (up to thousands of generations back) can be estimated by coalescence methods (Excoffier et al. 2021, Fedorca et al. 2024). In addition, the amount of linkage disequilibrium (LD) between genetic markers allows reliable estimates of contemporary N_e from a single set of sampled individuals (Waples 2024b, Santiago et al. 2024, 2025). Moreover, if high-density markers, such as SNPs, are available, LD between pairs of SNPs at different genetic distances allows the estimation of historical N_e for periods of time back to two hundred generations (Santiago et al. 2020, 2025). Thus, in this study we evaluated the current genetic diversity and the historical changes in N_e of one of the most important fisheries in Europe, the European hake, *Merluccius merluccius*.

European hake is an important fishery resource widely distributed across the eastern Atlantic and the Mediterranean Sea, being the third most valuable species in Europe. Fisheries stock assessments are performed yearly for the two Atlantic populations, the Southern and Northern stocks, and in different regions of the Mediterranean Sea. European hake stocks have decreased over several decades, with most of the populations severely affected in the late 20th century (ICES 2024; STECF 2025). As a consequence, several recovery plans and reductions in fishing effort were implemented, leading to the recovery in abundance of the two Atlantic stocks, which are now considered within safe biological limits (ICES 2024). The Mediterranean stocks, however, are still declining and severely overexploited (EU-STECF 2025). In spite of the well-known historical overexploitation rate of the majority of the fish stocks worldwide, few studies have focused on the loss of genetic diversity in exploited fish populations and their consequences for stock dynamics and sustainability, sometimes even with contradictory results (Pinsky and Palumbi 2014, Pinsky et al. 2021, Hoey et al. 2022). To fill this knowledge gap, we investigated genetic diversity and demographic changes in European hake stocks using a diverse set of population genomic tools, spanning allele frequency and linkage disequilibrium dynamics over time, as well as population genetic differentiation and relatedness.

Methods

Sampling

A total of 480 specimens of European hake were sampled in three different locations: 96 were collected at the North Sea (ATN), i.e. within the Atlantic Northern stock; another 96 specimens were collected on the Galician shelf in the Northwest Atlantic Iberian Peninsula (ATS), i.e. within the Atlantic Southern stock; and finally, 288 hake individuals were collected around the Balearic Islands in the Mediterranean Sea (MED), in the Geographical Sub Area (GSA) 5. All samples were collected by bottom trawl as follows: ATN during research survey No. 428 of FRV *Walther Herwig III* in summer 2019; ATS during the DREAMER-0617 research survey onboard R/V *Sarmiento de Gamboa* in June 2017 and from commercial fisheries between February and September 2017; MED during the MEDITS_ES05 2019 research survey, conducted between June 11 and 25 of 2019 onboard the R/V *Miguel Oliver*. For all specimens, length, sex and maturity stage were recorded, and fin clips were collected and stored in absolute ethanol for genomic studies (Suppl. Table S1). A map showing the ubication of sampling points within the three different locations is presented in Fig. 1.

Library preparation and sequencing

Genomic DNA of specimens was extracted from preserved fin clips using the DNeasy Blood & Tissue kit (Qiagen). The quality and concentration were checked using a Qubit fluorometer

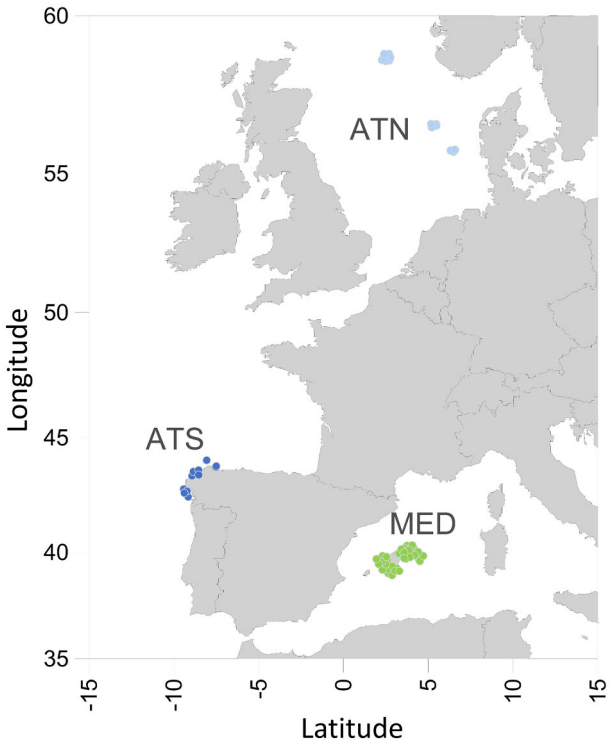

Figure 1. Location of samples collection. ATN: Atlantic Northern population. ATS: Atlantic Southern population. MED: Mediterranean population.

(Invitrogen) and a Nanodrop followed by gel electrophoresis prior to the preparation of single-digested RAD-seq libraries, following the protocol of Etter et al. (2011). Briefly, approximately 1 μg of DNA per sample was digested with the restriction endonuclease *SbfI* (CCTGCA|GG recognition site) (NEB) and ligated to modified Illumina P1 adapters containing 5 bp sample-specific barcodes to allow sample multiplexing. Pools of 48 individuals were sheared with a Q800R3 Sonicator (Qsonica) and size-selected for fragments spanning 300–500 bp via agarose gel electrophoresis. Purifications between steps were performed with QIAquick columns (Qiagen). Following the ligation of the Illumina P2 adapters, each genomic library was amplified using 12 PCR cycles and quality checked by a 2100 Bioanalyzer using a DNA 1000 chip (Agilent technologies). A step-by step protocol for library construction can be found at https://digital.csic.es/handle/10261/297272. Libraries were sequenced on a HiSeqX platform (Illumina) with 150-bp paired reads.

Data processing

The European hake reference genome was downloaded from NCBI (fMerMel1.1 Accession: GCA_964660975.1, Martinez et al. 2025). RADseq reads from the three European hake populations were demultiplexed using the process_radtags module from Stacks v2.0 (Catchen et al. 2011). The reads were subsequently processed for quality filtering with FastQC with default commands (Andrews 2010), and fastq files containing less than 500 000 reads were discarded. The final datasets, ATN (n = 94), ATS (n = 87), and MED (n = 269) were aligned to the reference genome using bwa-mem (Li 2013) with default parameters. Statistics on average number of reads, and genome coverage are presented in Suppl. Table S2 and Suppl. Fig. S1. Genomic alignments were then converted to bam format, indexed and sorted using samtools (Danecek et al. 2021).

Genetic diversity estimates

To get insight into the levels of genetic diversity and neutrality deviations, bam files for each population were separately analysed using ANGSD (Korneliussen et al. 2014) with the following commands: -doSaf, -minInd 50, -minMapQ 30, -minQ 20, -GL 1, with identical ancestral (-anc), and reference genome (-ref). To get the folded Site Frequency Spectrum (SFS), the realSFS script of ANGSD was used with the commands -fold 1, -maxIter 100. Statistics of the total number of analysed sites and number of SNPs in the SFS of each population are presented in Table 1. Genetic diversity estimates were obtained from the analysis of the SFS using realSFS saf2theta function and thetaStat do_stat scripts. In particular, the average amount of mutations within a population θ_W (Watterson 1975), and the average pairwise differences in the number of mutations within a population θ_π (Nei 1987) by site, were obtained. While the first considers only numbers of mutations, the second also considers their frequencies in the population. The difference between θ_W and θ_π is measured by the D statistic (Tajima 1989), which gives a direct insight into deviations from the neutral equilibrium model expected for populations with constant sizes across time, where the difference between the two genetic diversity estimates is expected to be null. Resampling of genetic diversity estimates was obtained by bootstrapping SFS in ANGSD with the script realSFS and commands: -boostrap 100, -fold 1. Each bootstrapped SFS was processed with saf2theta function and thetaStat do_stat scripts, returning estimates of genetic diversity for the 100 re-sampled SFSs for each population.

In order to compare genetic diversity estimates to those expected under a neutral model with constant size, 100 neutral SFSs were obtained with ANGSD by processing 100 coalescent simulations generated with ms software (Hudson 2002). Command lines were: ms 188 100 -t 35 for ATN, 174 100 -t 44 for ATS and 538 100 -t 49 for MED; where 188, 174 and 538 are the number of haplotypes in the sample, 100 is the number of replicates and -t is the composed estimate of genetic diversity $\theta \approx 4N_e\mu$, where μ is the mutation rate. θ was estimated for each population from θ_W obtained in the previous step for a genomic region of size 1000 bp (θ_W/total sites $\times$ 1000). Output files from ms were then processed with ANGSD the same way as for the real populations. Differences in levels of genetic diversity were tested by Wilcoxon rank sum test with continuity correction in R (R Core Team 2024). P-values were obtained by comparing Tajima's D of each population to the distribution of the 100 simulated dataset with ms for the same population. Finally, simulated SFS for two populations with constant size of N_e = 4 000 000 and N_e = 10 000 were obtained using ms and processed with ANGSD to constrast differences in the percentage of low frequency mutations between the expected and observed SFS (Suppl. Fig. S2).

Estimation of recent and ancestral effective population size by coalescence methods

The neutral theory of molecular evolution allows us to trace back in time the frequency of the currently ob-

Table 1. Statistics from the analysis of genetic diversity of European hake: θ_W (Watterson 1975), θ_π (Nei 1987), Tajima's D (Tajima 1989).

Population	Total analysed positions	Number estimated SNPs by SFS	Minimum analysed MAF	Percentage of SNPs MAF < 0.10 (Expected[a])	θ de Watterson (θ_W)	Pi (θ_π)	Tajima's *D*
ATN	15 678 479	3 219 523	$5.3e^{-3}$	91.6% (63.1%)	$1.9e^{-4}$	$8.9e^{-5}$	–1.85
ATS	16 390 231	4 221 403	$5.7e^{-3}$	94.2/% (61.3%)	$2.6e^{-4}$	$8.4e^{-5}$	–2.19
MED	19 970 377	6 795 686	$1.9e^{-3}$	94.7% (89.4%)	$8.7e^{-5}$	$3.2e^{-5}$	–1.92

[a]from a neutral model with constant population size and the same observed $\theta = 4N_e\mu$ as in the populations.
Note: Site Spectrum Frequency (SFS) was estimated using ANGSD software (Korneliussen et al. 2014, see Methods). ATN: Atlantic Northern population. ATS: Atlantic Southern population. MED: Mediterranean population.

served mutations by using the so-called coalescent methods. Demographic changes affect the Site Frequency Spectrum (SFS) with respect to that expected in a population evolving under constant population size. Thus, the observed SFS can be contrasted to different demographic models to get insight on the demographic history of the populations.

The full (i.e. invariant plus SNP positions) folded SFSs obtained in the genetic diversity analyses for each population by using ANGSD software (Korneliussen et al. 2014) were used for contrasting to the expectations of five demographic models by running coalescent simulations in FastSimCoal27 (*fsc27*, Excoffier et al. 2021): M0 = constant population size; M1 = one change in population size; and M2_1 = two changes in population size. Two additional models were ran for M2_1: 1) M2_2, by adjusting parameters search according to the results of M2_1 model, and 2) M2_3, by adding a population growth rate of 0.001 to M2_2 (see Suppl. Table S3 and Suppl. Fig. S3). Maximum Likelihood parameter estimates of contemporary (N_{ec}) and ancestral (N_{ea}) effective population size and the time of population changes (i.e. T1 and T2; see Suppl. Table S3 and Suppl. Fig. S3) were obtained by running 100 simulations for each model with commands -t file.pop.tpl (demographic model file, see Suppl. Table S3), -n 100 000 (number of coalescent simulations), -m (minimun allele frequency), -e file.pop.est (file with the search range of parameters to be estimated), -M (Maximum Likelihood) -L 40 (number of EMC cycles to search for the ML).

To obtain confidence intervals of the estimated parameters in *fsc27,* 100 boostrapped SFS were generated for each population by using the ANGSD script realSFS and the commands: -boostrap 100, -fold 1. Each bootstrapped SFS was processed as input in *fsc27* and simulations were ran using the -initvalues command and the .pv output file of the simulation ran with the best estimated parameters in the best fitted model (Suppl. Table S3, Suppl. Fig. S3). A total of 600 simulations were performed for each population, 500 for searching ML parameters and 100 for confidence intervals estimates of the parameters of the best fitted model (i.e. M2_3).

Additionally, ancestral effective population sizes of each population were investigated by non-model-based coalescent simulations as implemented in StairWay Plot software (*SWP*, Liu & Fu 2020). To run it we tested two different generation time lengths: 3 and 6 years. We used the full folded SFS (i.e. invariant plus SNP positions), a mutation rate of $2.5e^{-8}$, and 200 pseudoreplicates to estimate 95% CI of N_e. *SWP* reports the median and quartiles of N_e by years/generations.

Estimation of contemporary effective population size and recent historical demographic changes by linkage disequilibrium methods

Another way frequently used in population genetics to estimate the contemporary and recent changes in historical effective population sizes from a sample of individuals, is by looking at the pattern of linkage disequilibrium between markers (see, e.g. Fedorca et al. 2024). Here, we used currentNe2 and GONE2 (Santiago et al. 2025) to estimate contemporary and recent historical changes in N_e of the three populations of European hake.

First, we ran ANGSD software (Korneliussen et al. 2014) for each population with the following commands: -doPlink 2 (output in tfam tped format), -GL 1 (genotype likelihoods from samtools), -doMajorMinor 1 (major and minor allele frequencies from GL), -doMaf 2 (fixed Major, unknown Minor), -doPost (using frequencies as prior), -postcutoff 0.99 (set missing genotypes below cutoff), -SNP_pval $1e^{-6}$ (only positions with *p* value below $1e^{-6}$ were considered SNPs), -genomindepth 10 (minimum depth for consider a genotype). Output tfam and tped files were then transformed to plink map/ped format using PLINK v1.9 (Chang et al. 2015). The output map/ped files contained 3.34 millions of SNPs that were then filtered for a MAF of 0.01, resulting in 1 875 616 SNPs; 1 819 003 SNPs; and 2 050 864 SNPs for ATN, ATS, and MED, respectively. However, it has been shown that when working with RADseq, estimates of N_e by LD can be properly estimated avoiding high levels of missing data, which can substantially bias the estimates of both contemporary and historical N_e (Nunziata and Weisrock 2018, Marandel et al. 2020, Gargiulo et al. 2024). Thus, we restricted our estimates of N_e to a set of SNPs with no missing values. The resulting datasets consisted of 290 897, 211 177, and 287 144 SNPs for ATN, ATS, and MED, respectively.

Estimates of contemporary N_e by currentNe2 were obtained from pairs of SNPs from different chromosomes, which do not require assumptions on the recombination rate of the species. It was ran twice for each population assuming: 1) panmixia (default option) and 2) population structure with the -x option. If the estimates assuming structure are significantly larger than those assuming panmixia, this suggests population structure with limited migration among subpopulations. Otherwise, full panmixia or structure with very high migration rates among subpopulations can be inferred. The software provides 90% CI obtained by a neural network approach (Santiago et al. 2025).

GONE2 (Santiago et al. 2025) analyses were performed considering unknown phase, SNP pair windows with a maximum value of recombination frequency of 0.05 (default value), and an average recombination rate of 3.24 cM/Mb,

as obtained by Venu et al. (2024). This estimate is a reasonable one given the genome size (~700 Mb) and chromosome number (21) found for the European hake by Martinez et al. (2025). Given the uncertainty of the recombination rate, we estimated historical N_e using recombination values ranging from 3.0 to 3.5 cM/Mb and a 95% CI was obtained from the variation among these estimates.

Estimation of contemporary effective population size by the sibship method

An alternative estimate of the contemporary N_e can be obtained by inferring the number of full-sib and half-sib pairs in a sample of individuals using maximum likelihood from marker information. Estimates were obtained with the software COLONY (Jones & Wang 2010) assuming polygamy and the default options. A random sample of 20 000 SNPs were used for analyses to reduce computing time.

Estimation of average inbreeding and coancestry

Estimates of the coefficient of inbreeding (F) of individuals, and of the coancestry (relatedness or kinship, f) between pairs of individuals in each population, were obtained with the software EMIBD9 (Wang 2025). This method infers the allele frequencies of markers for a putative reference population of unrelated and non-inbred individuals with an Expectation-Maximization method, and then applies a maximum likelihood method to estimate the relatedness between individuals. Thus, the method provides estimates of inbreeding and coancestry of individuals (with range 0 to 1) relative to the mentioned reference population in the past.

Population genetic differentiation and admixture

We estimated different statistics to explore the genetic differentiation among the three stocks as a surrogate of the gene flow among them. The F_{ST}statistic (Wright 1951) measures the levels of genetic differentiation by quantifying the differences in allele frequencies among populations and it is considered closely related to the number of migrants (i.e. $F_{ST} \approx 1/[4N_em + 1]$, where m is the migration rate). Global estimates of F_{ST} among population pairs were obtained using ANGSD (Korneliussen et al. 2014) by processing site allele frequencies (i.e. .saf files) of all three populations together with the realSFS scripts: fst index, fst stats, and fst print.

To further explore genetic differentiation among populations we used ANGSD to get genotypes-likelihoods followed by NGStools ngsCovar script (Fumagalli et al. 2013) to analyse the covariance matrix of allele frequency in a Principal Component Analysis (PCA). To do this, we eliminated from the dataset the individuals with evidence of kinship among them ($n = 29$) and used the following commands of ANGSD: -minMapQ 30, -minQ 20, -minInd 421 (i.e. no missing data), -doCounts 1, -GL 1, -doMajorMinor 1, -doMaf 1, -SNP_pval 1e^{-6}, -doGeno 32, -doPost 1; which resulted in 553 666 SNPs. Next, ngsCovar script was run with the commands -nind 421 (i.e. no missing individuals), -call 1 (i.e. perform genotype calling), and -minmaf 0.01; which resulted in 212 562 SNPs for all three populations altogether. A PCA plot was obtained by using the Rscript of NGStools ngsPopGen (Fumagalli et al. 2013).

We inferred population structure and admixture proportions between populations by running ANGSD and NGSadmix (Skotte et al. 2013) with beagle genotype likelihoods with the following commands of ANGSD for all population together -minMapQ 20, -minQ 20, -doCounts 1, -GL 2, -doMajorMinor 4, -doMaf 1, -doGlf 2, -SNP_pval 1e^{-6}, and -minind 421 (i.e. no missing data). Then, NGSadmix was run for 3 alternative clustering of individuals (K = 2 to 4), The evaluation of the best clustering model (K) was performed using EvalAdmix software (Garcia-Erill and Albrechtsen 2020).

Finally, as linkage among SNPs could obscure genetic differentiation and structure analyses we obtained a reduced dataset of SNPs randomly distributed across the genome. This dataset was produced by filtering SNPs by MAF > 0.01 and minimum depth of 10 reads using vcftools (Danecek et al. 2021). The resulting vcf was filtered by taking one random SNP for each alignment block. The final dataset consisted of 16 982 SNPs, which was analysed using alternative methods. F_{ST} indices were obtained by using vcftools (Danecek et al. 2021). Population structure and admixture proportions were estimated using ADMIXTURE v1.3.0 (Alexander et al. 2009), assessing two to four genetic clusters (K = 2–4). Finally, a PCA analysis was performed using PLINK v1.9 (Chang et al. 2015). R (R Core Team 2024) was used to plot all results. Additionally, for this reduced SNP dataset, the phylogenetic relationships among all individuals were estimated by computing a Maximum Likelihood (ML) tree with RAxML-NG (Kozlov et al. 2022) using the unphased diploid genotype model (GT10) described in Kozlov et al. (2022) and 100 bootstrap replicates. Results were visualized using FigTree v1.4.4 (Rambaut 2018).

Correlation between N_e and stock parameters

Three stock parameters were used to investigate the relationship between the historical estimates of N_e across time as obtained by GONE2 and stock trends. First, recruitment (REC, measured in number of individuals) is the result of the reproductive success of the population and hence clearly linked to N_e. Second, in fisheries management the spawning stock biomass (SSB, measured in tonnes) is normally used as an index of the adult population contributing to the annual recruitment, i.e. potentially the gene pool of the next generation. And third, fishing is the major source of mortality and hence potentially impacting N_e; we used here the reported catch (measured in tonnes) as used in the stock assessments as a measure of mortality. Estimates of stock parameters were obtained from the 2020 stock assessments by GFCM in the GSA5 section for Balearic Islands (https://www.fao.org/), and ICES (2024) for the two Atlantic stocks, the Northern and Southern stocks. Spearman rank correlations were performed between the three stock parameters and historical estimates of N_e using R (R Core Team 2024). Since N_e is estimated per generation whereas stock parameters are estimated annually, to correlate stock parameters to N_e per generation, the year of sampling collection (e.g. 2019) and the two previous years were considered to have the N_e of generation 1, the period between 2016 to 2014 the N_e of generation 2, the period between 2013 to 2011 the N_e of generation 3, etc. The three years interval was selected as generation time based on the estimated age at which 50% of individuals become mature for sexes-combined and the mean age of mature population (Piñeiro and Sainza 2003). For a better visualization of the joint historical trends in N_e and stock parameters, the latter were averaged over three-year intervals, beginning with the year of sampling (i.e. 2017 or 2019) and proceeding backwards.

Additionally, as annual adult abundances (N_a) for the hake Southern stock were available from the ICES WGBIE, the relationship between N_a and N_e was also explored for this population. This ratio is important for assessing the relative risks that demographic, environmental, and genetic factors may pose to population persistence, especially since N_e is typically much smaller than N_a in natural populations (Palstra and Fraser 2012, Hoban et al. 2020).

Results

Decline in genetic diversity and effective population size of European hake

The levels and pattern of genetic diversity in the European hake populations were analysed by using four different population genetics methods: **1**) deviations from neutral expectations, **2**) coalescent simulations, **3**) genome-wide linkage-disequilibrium estimates, and **4**) sibship analyses.

1. Current levels of genetic diversity and deviations from neutral equilibrium

The two estimators of genetic diversity (i.e. $\theta \approx 4N_e\mu$, Nei 1987) of European hake showed differences among populations, with ATS population showing significantly higher values of the average amount of allelic variants within a population (Watterson´s θ_W), and MED significantly lower values of the average pairwise differences in the number of variants within a population, θ_π, (Wilconson test correct p-value < 0.001 for both genetic diversity estimates, Fig. 2 A-B). Tajima's D showed negative values in all three populations (Fig. 2 C), significantly deviating from the expectations of constant population size obtained by coalescent simulations ($p < 0.01$). Moreover, the percentage of SNPs with minimum allele frequency lower than 10% was significantly higher than expected under a neutral model of constant population size (2-sample test for equality of proportions, p-value $< 2e^{-9}$). Summary statistics are presented in Table 1. Suppl. Fig. S2 shows differences in the proportion of SNPs with minimum allele frequency lower than 10% between the observed in populations and the expected under constant population size with a contemporary $N_e = 10\,000$ or $N_e = 4\,000\,000$.

2. Recent and ancestral effective population size estimated using coalescent simulations

FastSimCoal27 (Excoffier et al. 2021) was used to simulate five different demographic models for each population (Suppl. Fig. S3 and Suppl. Table S3). M0: constant population size; M1: one change in population size, and M2_1: two population changes. Additionally, two refined models, aimed to explore more complex demographic scenarios, were developed to improve model fit based on M2_1: 1) M2_2: adjusted model parameters based on the results of M2_1, and 2) M2_3: further refined M2_2 by incorporating recent population growth (see Suppl. Fig. S3, and Suppl. Table S3). The results show significant differences between M0 vs M1, and M1 vs M2 models (AIC > 4), but not significant differences among M2 models. However, M2_3 got a slightly better Maximum Likelihood adjustment than the other models (Suppl. Fig. S3 and Suppl. Fig. S4), and therefore we took the parameters estimated by this model as a reference. Confidence intervals of the contemporary (N_{ec}) and ancestral (N_{ea}) effective population sizes are presented in Table 2. Overall, coalescent simulations indicate that, in terms of long-term genetic variation, all three populations share a common history with large ancestral estimates and sharp reductions in recent effective population sizes. Estimates of contemporary N_{ec} from *fsc27* were of the order of a few thousand individuals with confidence intervals ranging from a few hundred to hundred thousands (Table 2). In contrast, average estimates of ancestral N_{ea} where higher than 2 million individuals ranging from 100 000 to more than seven millions (Table 2). Two times of demographic changes (i.e. coalescent times T1 and T2, see Suppl. Fig. S3) were found. The most recent (T1) was found within the last 5 generations, evidencing a recent decline, and the oldest T2 around 200 000 generations ago, showing an ancestral expansion. Confidence intervals for coalescent times were narrow for T_1 (i.e. population bottleneck), from 2 to 5 generations ago, and varied from 193 000 to 324 000 generations ago for T_2 (i.e. population expansion).

Non-model-based coalescent simulations as implemented in StairWay Plot2 (*SWP*, Liu & Fu 2020) showed large effective population sizes around 2 million individuals in all three populations. This large ancestral effective population size appears to have been nearly constant since nearly 600 000 to 2 000 years ago with a sharp decline beginning in recent times, < 1 000 generations ago (Table 2; Fig. 3-A). Similar results were found for a generation time of six years. Thus, model-based and non-model based coalescent simulations show the same historical trend in all three populations with ancestral large and reduced contemporary N_e. Estimates from *fsc27* showed wider confidence intervals than those from *SWP*, likely because the former are based on bootstrapping SFS, while the latter are based on pseudo-boostrapping. The mean (and confidence intervals) of the percentage of population declines estimated as 100-($N_{ec}/N_{ea} \times 100$) were 99% (97–99) for ATN, 90% (38–99) for ATS, and 99% (87–99) for MED. For *SWP* all estimates were > 99%.

3. Estimates of effective population sizes from linkage disequilibrium between SNPs and the sibship method

Results from currentNe2 (Santiago et al. 2025) showed almost identical estimates of contemporary N_e for the ATN (9 485) and ATS (9 498) populations, and the lowest one for the MED population (3 314), which was significantly lower than those from the Atlantic populations (Table 2). Running the program with the option -x to assume population structure gave the following estimates of N_e (Confidence Intervals): For ATN, 5 628 (3 686–8 593); for ATS, 10 413 (5 870–18 471); and for MED, 1 870 (1 674–2 089). These three estimates are not significantly higher than those assuming panmixia in Table 2, indicating that none of the three populations shows signs of population structuring with restricted migration from other populations.

The sibship method, which has less power than the linkage disequilibrium method when the values of N_e are very large, provided an infinite estimate for ATS, an estimate of 9 012 for ATN, with an infinite upper limit, and a much lower estimate of 1 670 for the MED population. The lowest estimate of N_e for the MED population is fully concordant with the estimates of the average inbreeding coefficient (F) and coancestry coefficients (f), being the highest for the MED population (average F of 0.22 *vs* about 0.02 for the Atlantic populations; and average f of 0.28 *vs* < 0.03 for the Atlantic populations; Suppl. Table S4). Additionally, and also in agreement with the above results, the software COLONY did not infer any full-

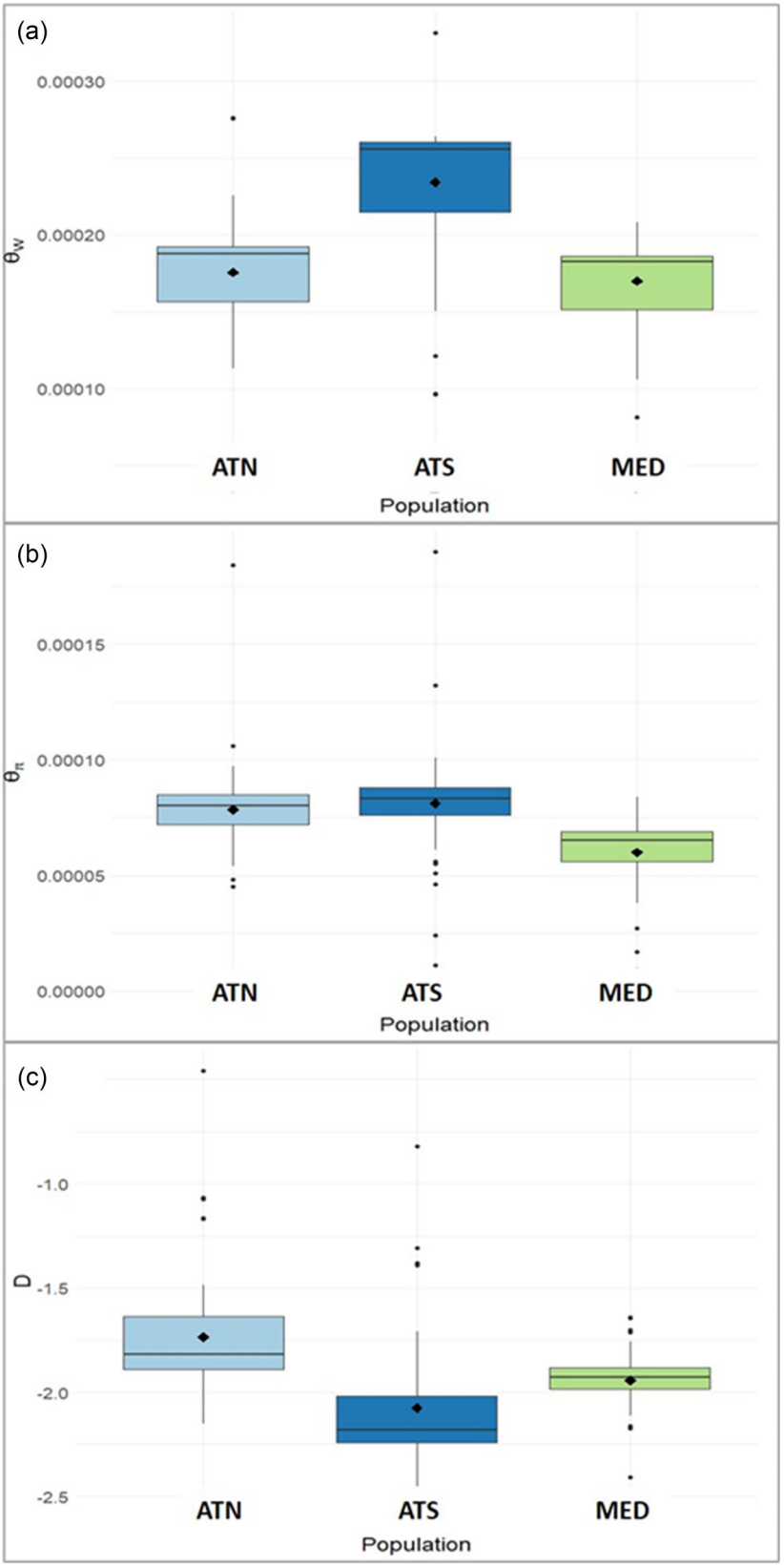

Figure 2. Differences in genomics levels of genetic diversity in European hake populations ATN (light blue): Atlantic Northern population. ATS (dark blue): Atlantic Southern population. MED (green): Mediterranean population. **(A)** θ_W by site, **(B)** θ_π by site, and **(C)** Tajima's *D*. Boxplots represent the variation in genetic diversity estimates among 100 replicates obtained by bootstrapping the Site Frequency Spectrum (SFS) of the populations. Differences among pairwise populations are significant (Wilconson test correct *P*-value $2e^{-16}$). Diamonds = mean values, dots = outlier values, horizontal line = median values, box = 1st and 3rd quartiles, vertical lines = lower and upper limits.

Table 2. Estimates of contemporary (N_{ec}) and ancestral (N_{ea}) effective population size from the coalescent simulations FastSimCoal27 (fsc27, Excoffier et al. 2021) and StairWay Plot (SWP, Liu & Fu 2020, linkage disequilibrium between SNPs (currentNe2 and GONE2, Santiago et al. 2025), and sibship method (COLONY, Jones & Wang 2010).

	fsc27		SWP		GONE2			
	N_{ec}	N_{ea}[a]	N_{ec}[b]	N_{ea}[c]	N_{ec}	N_{ea}[d]	currentNe2 N_{ec}	COLONY N_{ec}
ATN	4 366 (1 150–1.5e[d])	2.0e^7 (3.0e^5–1.5e^9)	4 397 (2 024–17 672)	2.0e^6 (1.5e^6–4.0e^6)	3 500 (2 878–4 122)	4.3e^6 (3.7e^6–4.8e^6)	9 485 (4 802–18 736)	9 012 (3 881–∞)
ATS	13 719 (8 561–4.0e^5)	5.0e^6 (1.0e^5–7.0e^6)	10 743 (1 661–61 541)	2.0e^6 (1.5e^6–4.0e^6)	10 294 (8 581–12 007)	1.5e^6 (1.2e^6–1.8e^6)	9 498 (4 588–19 665)	∞ (1 —∞)
MED	12 153 (312–2.4e^5)	8.0e^6 (1.8e^5–3.5e^9)	1 851 (370–3 832)	2.0e^6 (1.5e^6–4.0e^6)	548 (382–714)	2.2e^5 (4.2e4[d] –4.0e^5)	3 314 (2 753–3 988)	1670 (1 368–2 104)

[a]. at the coalescent time T_2: 225 079 (ATN), 279 427 (ATS), 277 432 (MED) generations ago. See Suppl. Table S3.
[b]. at 10 generations ago.
[c]. at about 1 000 to 200 000 generations ago (See Fig. 2-A).
[d]. at 25 generations ago (See Fig. 2-B).
Note: For fsc27 the best likelihood estimates of N_{ec} and N_{ea}, and 1st and 3rd quartiles of these estimations obtained by 100 re-sampling of the Site Frequency Spectrum (SFS) of each population are presented. Confidence intervals for GONE2 estimates were obtained from variation across estimates assuming slightly different recombination rates (see Methods and Suppl. Fig. S5). For SWP, currentNe2, and COLONY presented confidence intervals are as given for each program (see Methods). ATN: Atlantic Northern population. ATS: Atlantic Southern population. MED: Mediterranean population.

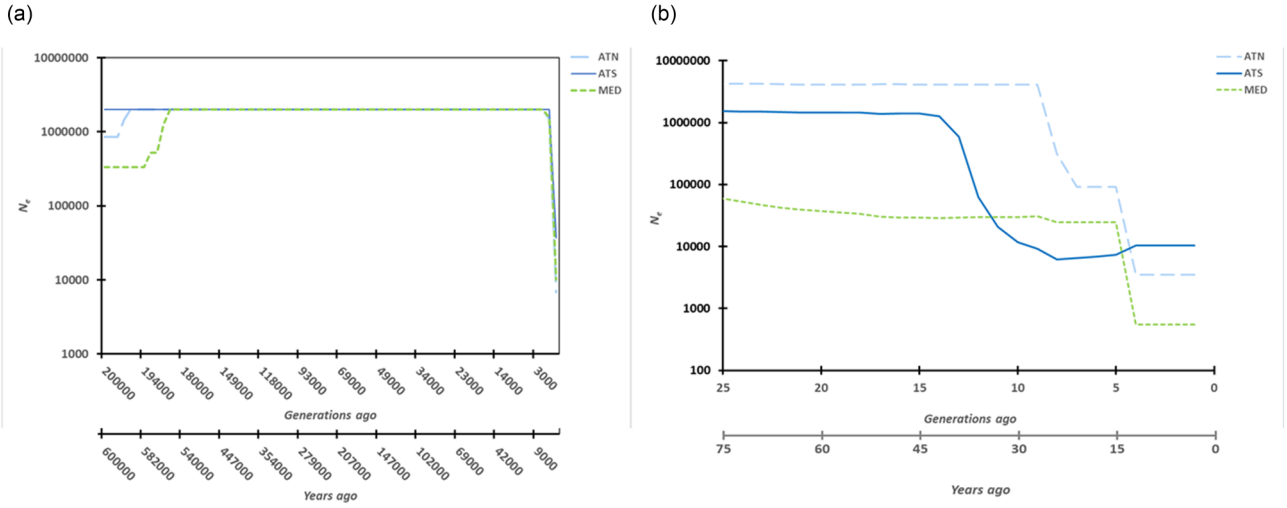

Figure 3. Long and short-term demographic history of European hake populations, **(A)** Estimated ancestral demographic history using StairWay Plot (Liu & Fu 2020) with a generation time length of 3 years. Graph was constructed by representing means values of N_e in periods of 3 000 years (i.e. 1 000 generations). Folded full (i.e. invariant and SNPs positions) Site Frequency Spectrum (SFS) were obtained with ANGSD (Korneliussen et al. 2014) for each population (see Methods). Mutation rate = 2.5e^{-8}. **(B)** Recent historical effective population sizes (N_e) of European hake populations estimated by GONE2 (Santiago et al. 2025). Generations are scaled to years for a generation time length of 3 years. Results from *fsc27* (Excoffier et al. 2021) support both ancestral and recent demographic history (see Table 2). ATN: Atlantic Northern population. ATS: Atlantic Southern population. MED: Mediterranean population.

sib or half-sib family for the ATS population (explaining the infinite estimate of N_e for that population), and inferred no full-sib families and two half-sib families of two members each for the ATN population. In contrast, for the MED population, three full-sib families were inferred with 8, 3 and 2 members, respectively, and 10 half-sib families (one with 5 half-sibs, one with 4 half-sibs, four with 3 half-sibs, and four with 2 half-sibs), reflecting the much lower estimate of N_e for this population.

Historical estimates of N_e from GONE2 indicated a decline of at least two orders of magnitude across all three populations over the past 25 generations (Table 2, Fig. 3-B, and Suppl. Fig. S5). Again, in agreement with the above results, the contemporary estimate (N_{ec}) for the MED population was significanty lower than those for the Atlantic populations, and that for ATS significantly higher than that for ATN (Table 2). The declines observed by GONE2 in the two Atlantic populations seemed to have occurred somewhat earlier in the ATS population (approximately 12 generation back) than in the ATN population (about 8 generations back), with some recent recovery in the ATS population (Fig. 3-B, Suppl. Table S5). The Mediterranean population, in contrast, seems to have had a much lower ancestral effective size sharply dropping again in the most recent generations. However, the time of decline for this population and the ancestral estimates varied substantially depending on the recombination rate assumed (Suppl. Fig. S5).

Thus, similarly to results obtained by coalescent simulations, estimates of historical N_e by linkage-disequilibrium indicate a relatively recent strong decline in all three populations. This decline amounts to more than 95% for all three populations estimates.

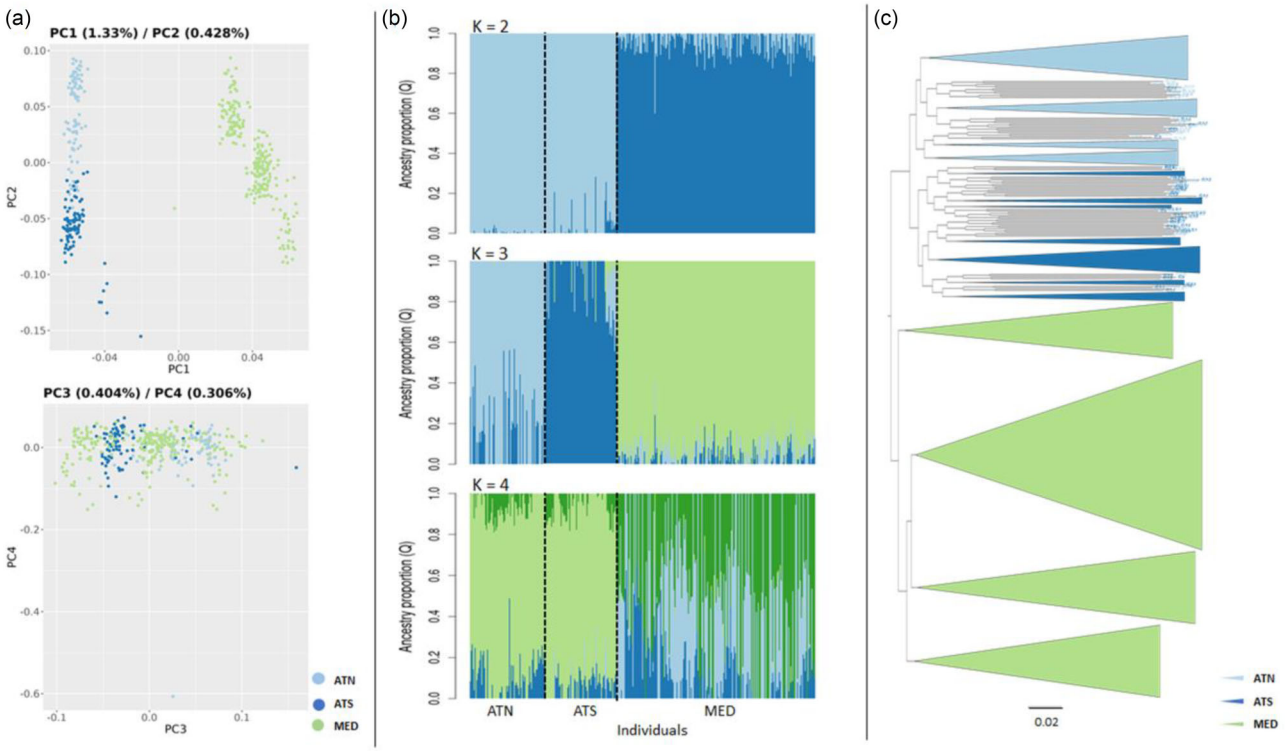

Figure 4. Genetic differentiation among European hake populations. **(A)** Distribution of individuals based on the first two components of the principal component analysis (upper graph) and third and fourth PCA components (lower graph), estimated using the full SNPs dataset (212 562 SNPs, see Methods). **(B)** Population structure plots showing the ancestry of each individual (vertical bars), estimated using the same SNP dataset as for PCA. Models assumed K = 2, 3 and 4 genetic clusters (from top to bottom), with K = 2 being enough to explain the differences among populations the according to EvalAdmix analyses (Garcia-Erill & Albrechtsen 2020) (see Suppl. Fig. S8). **(C)** Simplified Maximum Likelihood tree between the 450 hake individuals as inferred by the concatenated SNP matrix by raxML-NG using the reduced dataset of 1 SNP by RAD locus (16 982 SNPs, see Methods). Principal Component Analyses and Admixture analyses for this reduced SNP dataset are presented in Suppl. Fig. S7. ATN: Atlantic Northern population. ATS: Atlantic Southern population. MED: Mediterranean population.

Population genetic differentiation and admixture of European hake populations

Global genomic estimates of genetic differentiation measured by F_{ST} index were not significantly different between population pairs (Suppl. Fig. S6), with mean values of 0.007 for ATN-ATS pair, and 0.010 for both MED-Atlantic pairs. Similar results were obtained by using the reduced set of unlinked SNPs, although slightly higher F_{ST}estimates were found between the Atlantic populations and the Mediterranean one (Suppl. Table S6). Principal Component Analysis (PCA) show a clear differentiation between Mediterranean and Atlantic populations (Fig. 4-A), although first and second components only explain around 2% of the total genetic variation. Third and fourth component explained less than 1% of the variation (Fig. 4-A). Similar results were found by using a reduced set of unlinked SNPs but, like F_{ST} estimates, the percentage of variation explained by PC1-2 was also higher (Suppl. Fig. S7-A).

The results of the admixture analyses are presented in Fig. 4-B. EvalAdmix (Garcia-Erill & Albrechtsen 2020) estimates indicate that two genetic clusters are enough to explain the observed pattern (K = 2, Atlantic vs Mediterranean populations; Suppl. Fig. S8). Similar results were found when using the reduced set of unlinked SNPs (Suppl. Fig. S7-B).

The phylogenetic tree inferred for all individuals shows as well a deepest (and slightly supported) split defining two main clades formed exclusively by MED vs. Atlantic populations. In the Atlantic clade, two further subclades can be noted; one with most ATN and the other with most ATS individuals. However, some individuals of these latter clades appear intermingled (Fig. 4-C).

Relationship between stock parameters and N_e in European hake populations

The temporal trends of the estimated stock parameters for each population (Suppl. Table S7) and the historical N_e obtained using GONE2 for a recombination rate of 3.24 (Suppl. Table S5) are shown in Fig. 5. The length of the time series for stock parameters varied among stocks, with the first year of available data being 1978 for ATN, 1960 for ATS, and 1980 for MED (Suppl. Table S7). Comparable trends in all stock parameters were observed for the ATS population, characterised by a pronounced decline in all variables during approximately the same period (1980–2000). This pattern is further evidenced by the positive and statistically significant correlation among historical N_e and stock parameters obtained for this stock (Table 3 and Suppl. Fig. S9). For ATS, SSB and N_e exhibit a highly similar pattern ($r = 0.96$), but, interestingly, the decline in SSB during the 1980s preceded the decline in N_e (Fig. 5). The ratio between the effective population size and the annual adult abundance N_e/N_a (Fig. 6) remained relatively constant for several generations (approximately 1960–1980),

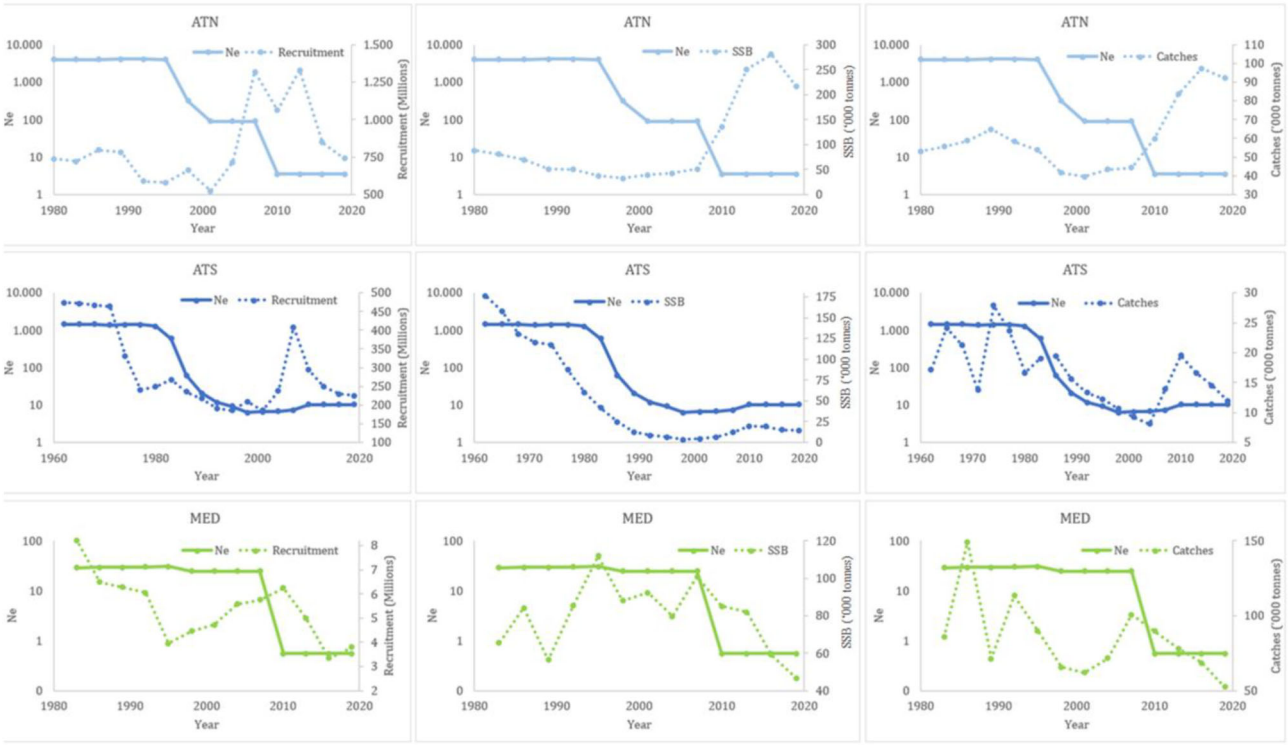

Figure 5. Relationship between the historical recent N_e and stock parameters: recruitment (first column), spawning stock biomass (SSB, second column) and catches (third column) for the Atlantic Northern, ATN, Atlantic Southern, ATS; and Mediterranean, MED stocks. Note the differences in the starting time of the abundance data series. N_e estimates are shown in log scale and stock parameters in real scale for a better visualisation of the trends. Spearman's rank correlations for non-transformed data are presented in Table 3. ATN: Atlantic Northern population. ATS: Atlantic Southern population. MED: Mediterranean population.

Table 3. Spearman rank correlations between stock parameters and effective population size by years.

	Recruitment	SSB	Catch
ATN	−0.24 (0.12)	**-0.49** (0.35)	−0.20 (**0.84**)
ATS	**0.58**	**0.96**	**0.73**
MED	0.23	0.30	0.30

Note: Significant correlations ($P < 0.001$) are shown in bold. Values in parenthesis for the ATN population show correlations for the period from 1978 to 2006. Scatterplots of the correlations among historical N_e and SSB are presented in Suppl. Fig. S9. ATN: Atlantic Northern population. ATS: Atlantic Southern population. MED: Mediterranean population.

followed by a steady decline over the next two decades (1980–2010), decreasing by nearly 20-fold. Since then, the ratio has stabilized, suggesting a shift to a new demographic regime.

For the MED population, for which historical abundance data are more limited, stock parameters exhibited substantial fluctuations, whereas N_e remained largely constant during 1980–2007. Consequently, no significant correlations were detected (Table 3 and Suppl. Fig. S9). It is noteworthy, however, that the sharp decline in N_e observed between 2007 and 2010 was followed by subsequent decreases in recruitment, SSB, and, correspondingly, in catches, remaining all at low level since then (Fig. 5).

In the ATN population, trends in stock parameters and N_e appeared to follow opposite patterns, with stock levels increasing markedly in recent years while N_e declined sharply (Fig. 5). This relationship was reflected by negative correlations, significant only for SSB (Table 3 and Suppl. Fig. S9). However, when the most recent years were excluded, the correlations became positive, although significant only for catch (Table 3 and Suppl. Fig. S9).

Discussion

The patterns of genomic diversity across the European hake populations were found to significantly deviate from neutral expectations of constant population sizes. The significant skew of the Site Frequency Spectrum towards low-frequency variants revealed by the Tajima's *D* statistic is compatible with the demographic scenery of a recent population bottleneck (Tajima 1989). Consequently, linkage disequilibrium analysis of historical N_e as implemented in GONE2 (Santiago et al. 2025) and model-based coalescent simulations as implemented in FastSimCoal (*fsc27*, Excoffier et al. 2021) provide evidence of a recent sharp decline in N_e since around 2–15 generations ago. Moreover, both methods infer a large ancestral effective population size of more than 100 fold the estimated contemporary N_e. Additionally, non-model-based coalescent simulations as implemented in StairWay Plot2 (Liu and Fu 2020) also support large ancestral effective population size of around 2 million individuals between nearly 200 000 to 500 generations ago, and a sharp decline in N_e since then.

Estimates of contemporary $N_e$obtained by currentNe2 (Santiago et al. 2025) and COLONY (Jones & Wang 2010) are within the same order of magnitude as those from *fsc27* and GONE2. All estimates suggest that the Mediterranean population has a significantly lower contemporary effective size than the Atlantic populations, and the Atlantic Southern population a slightly higher $N_e$than the Atlantic Northern.

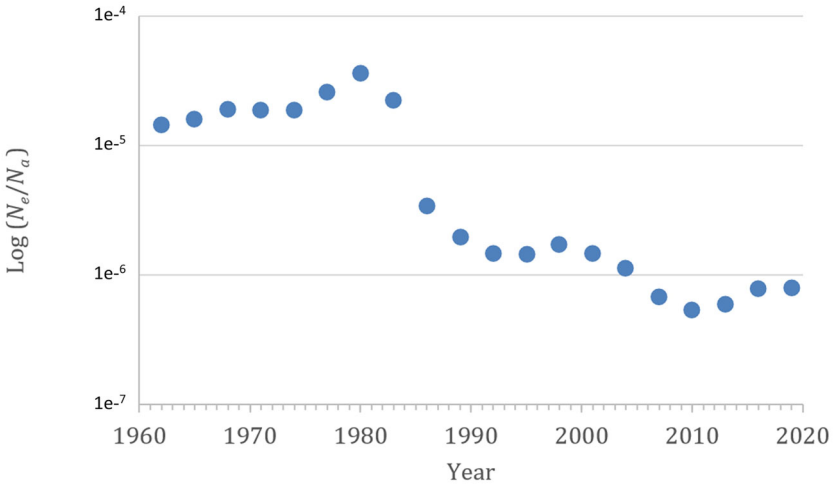

Figure 6. Ratio between the recent historical effective population size (N_e) as obtained by GONE2 and the annual adult abundance (N_a) for the Atlantic Southern population (ATS).

These results are also supported by a considerably higher average inbreeding coefficient and average coancestry coefficient between individuals for the Mediterranean population, where three full-sib families and ten half-sib families were detected.

The decline in N_e detected in the Atlantic Southern stock is also supported by previous studies by Pita et al. (2017) and Fernández-Míguez et al. (2023) who also found a recent decline in the effective population size of nearly 40-fold its size of 1978. Although this reduction in population size is lower than that we found for the same stock, reported estimates of N_e were also lower than those estimated here. This discrepancy in the magnitude of N_e estimates is possibly due to the enhanced power of our analyses, as we used hundreds of thousands of SNPs along with methods allowing trace back N_e estimates for a longer period. In this regard, using genotype-likelihood-based methods as implemented in ANGSD (Korneliussen et al. 2014) allows working at the minimum allele frequency possible according to sample size and alignments quality, likely improving the statistical power of both, coalescent and disequilibrium-based estimates of recent and historical effective sizes.

In summary, all different approaches used to study the demographic history of the European hake populations, coincide in the same general pattern of large ancestral effective population size and a sharp decline in recent times. However, recent times could have a wide interval in evolutionary periods that may include from few to thousands of generations ago. Nonetheless, both *fsc27* and GONE2 analyses show evidence that strong declines ocurred within the last 15 generations (i.e. nearly 50 years ago), although from our analyses it is not completely clear when the main decline of the Mediterranean population took place. Likelihood differences among the *fsc*27 results from the three models with two population changes were not significantly different among them, suggesting that demographic recent history of this population could include more than one bottleneck. In fact, GONE2 results show a considerable variation in the time of decline depending of subtle changes in the recombination rate used to obtain the estimates (Suppl. Fig. S5). Thus, it should be highlighted that variation in parameters unknown for the species, such as mutation and recombination rates, could produce changes in the estimated time and magnitude of the decline.

In addition, it must be considered that any method used to estimate N_e has limitations related to its theoretical assumptions and methodological constraints. To highlight, estimates of historical N_e are more precise when a refined genetic map is available (Santiago et al. 2020). This is usually unavailable for non-model species, and an average recombination rate constant across the genome must be assumed instead. We assumed an average recombination rate found for fish species and allowed for a certain variation in this parameter, which gave slightly variable outcomes particularly for the Mediterranean population. Moreover, one assumption of the estimates of N_e is that generations are discrete. Overlapping generations usually imply a downward bias of about 10% to 50% of estimates with the LD method (Waples 2024b). Computer simulations performed by Santiago et al. (2020) assuming three cohorts per generation involved downward biases of the estimated recent N_e by GONE of about 40%. Thus, the recent estimates of GONE2 in Table 2 could be underestimates by that approximate proportion. Another assumption of all methods of estimation of N_e is that the population under study is a closed one without migration from other populations. Estimates of contemporary N_e can be biased when the population is structured with migration between subpopulations (Waples 2024b). Our results from currentNe2, however, show that estimates assuming population structure are not significantly larger than estimates assuming panmixia, which suggests that population structure with limited migration between subpopulations is not the most likely scenario for the analyzed European hake populations. Finally, estimating N_e in large marine populations is generally difficult, and it has been concluded that sample sizes of around 1% of the total number of individuals are necessary to obtain accurate estimates (Marandel et al. 2018). Thus, as recent N_e values estimated by *fsc27*, GONE2, and currentNe2 in the three hake populations are generally below or around 10 000 individuals (Table 2), the analysed sample sizes in this study (i.e. nearly 100 for the Atlantic populations and more than 200 for the Mediterranean one) seem to be in the limit to obtain robust recent N_e estimates.

The positive correlation between historical N_e and stock parameters is particularly notable in the Atlantic Southern stock. The high fishing pressure observed prior to 1995 appears to explain the sharp decline in N_e. Interestingly,

spawning stock biomass (SSB) and N_e were highly correlated ($r = 0.96$) although the effects of intense fishing pressure were first evident in SSB before becoming apparent in N_e. This suggests that the stock was initially able to withstand a moderate reduction in SSB. However, continued declines in SSB ultimately may have led to a reduction in the stock's reproductive potential, an erosion of reproductive resilience (Lowerre-Barbieri et al. 2017), and a subsequent rapid decline in N_e. Despite the reduction in fishing pressure over the past two decades, the population has not recovered their ancestral levels of genetic diversity. SSB remains approximately ten times lower than in the 1960s, likely due to the persistently low N_e (Hoey et al. 2022). Notably, the ratio between N_e and the annual adult abundance (N_a) is now much lower than it was 20 generations ago (circa 1960), indicating a decoupled recovery dynamic between demographic and genetic indicators. While adult abundance has returned to levels observed in the mid-1980s, N_e remains low. This disparity likely explains the lack of full recovery of the Atlantic Southern stock, despite the implementation of a recovery plan by the European Commission in 2006 (Reg. EC No. 2166/2005). Positive correlations were also observed in the Mediterranean population, although they were weaker and not statistically significant, likely due to the shorter historical time series of stock data and the limited variation in N_e during the observed period, as the majority of the decline in N_e seems to have occurred earlier. Finally, in the case of the Atlantic Northern stock, where negative correlations were observed for the whole data series, it is important to consider that the genomic data analyzed in this study were collected in the North Sea. Historically, hake biomass, and consequently, catches, in the North Sea have been low (Staby et al. 2018), whereas the majority of the population and fishery activity has been concentrated in the Celtic Seas and the Bay of Biscay. As a result, historical stock assessments conducted by ICES were largely based on data (from surveys and catches) collected outside the North Sea. Thus, the stock parameters used in this study could reflect that geographic bias. Although the North Sea is officially considered part of the Atlantic Northern stock, the lack of positive correlation between stocks parameters and N_e suggests that hake in this area may have followed a different population trajectory than the main part of the stock. Whether this implies that the North Sea population constitutes a separate stock remains a subject of ongoing debate.

Previous studies on the genetic structure of European hake in the Atlantic have reported differentiation between the North Sea (or Norwegian Sea) and other Atlantic regions (Milano et al. 2014, Westgaard et al. 2017, Leone et al. 2019), though the genetic boundary among them is not clear. The present study found lower levels of genetic differentiation between individuals from the North Sea and the Galician shelf (part of the Southern stock) than those found by Leone et al. (2019). However, it is to note that the North Sea population analysed here is located south of that analysed by Leone et al. (2019). Our results are in agreement with earlier studies by Pita et al. (2014) using mitochondrial and microsattelites markers, which suggest significant migration rates between North and South Atlantic stocks. Actually, it has been proposed that the spawning biomass of the Atlantic Northern stock could play a crucial role by allowing recovery of the Southern stock (Pita et al. 2014, 2017).

The results of EvalAdmix (Garcia-Erill & Albrechtsen 2020) indicated that clustering our data in two groups is enough to explain the observed differences, supporting the low levels of genetic differentiation found by F_{ST}indices, PCA, admixture, and phylogenetic analyses among the two Atlantic populations. Gandra et al. (2021) found indices of genetic differentiation negatively correlated with the fish's economic importance. These authors suggest that the lack of genetic differentiation in abundant intensively fished species is due to "genetic homogenization driven by fisheries-induced bottlenecks and subsequent population replacement through immigration" (Gandra et al. 2021). The low genetic differentiation found between the two Atlantic populations and the evidence of recent bottlenecks for both would support the hypothesis of Gandra et al. (2019). Nevertheless, it is clear that further investigation is needed, incorporating genomic and demographic data from additional Atlantic locations.

Although the genetic differentiation between the Mediterranean and the Atlantic populations is slightly higher than between the two Atlantic populations, it is still very low, suggesting that despite the Gibraltar Straight and the Almería-Oran barrier there is great connectivity between them. This high gene flow could be explained by the almost continuous distribution of the species through the Mediterranean and the Atlantic coasts. Nonetheless, the more restricted genetic flow among Mediterranean and Atlantic populations could have consequences in the recovery of Mediterranean populations when gene flow is additionally diminished by decays in N_e(Hauser & Carvalho 2008).

Finally, the Atlantic Southern population shows signs of N_e recovery acompanying the recovery of abundances, although more slowly. Thus, if the observed recovery tendency in N_e and abundance are maintained, it could be expected that the remaining standing genetic variation will increase in allele frequency, while the population will gain new genetic variation by mutation. Altogether, this would reinforce the adaptive potential of the Atlantic Southern population, favouring also neighbouring populations. Thus, European hake populations could widely benefit of the implementation of spatial fisheries management considering meta-population dynamics in the stock assessments (Fogarty & Botsford 2007).

Although the primary goal of stock assessments is not genetic recovery, but abundance, broader management objectives increasingly include the maintenance of population structure, genetic diversity, and therefore, the long-term adaptive potential, especially under environmental changes. Besides the uncertainties regarding the exact time of N_e declines and the length of generation intervals, the overall results presented in this study indicate that the integration of population genetic estimates into stock assessment is not only possible, but crucial for optimal long-term sustainable fisheries management.

Conclusions

Using four different population genetic approaches, our study provides evidence of a recent and substantial decline in genetic diversity and effective population sizes in European hake populations. These declines have reduced N_e by at least two orders of magnitude compared to ancestral levels and are consistent with documented episodes of overexploitation. Such drastic reductions in N_e can have profound consequences for the evolutionary trajectory of hake populations, potentially disrupting gene flow patterns and reducing adaptive capacity to respond to environmental changes. Nonetheless, a likely recent

small increase in effective size was observed in the Atlantic Southern population, suggesting that management strategies implemented in recent years may have begun to support the recovery of the stock's genetic integrity although at very slow pace. Finally, the detailed methodology implemented in this study could be broadly applied in fisheries science to forecast the evolutionary pathways of other exploited marine species.

Acknowledgements

The authors are grateful to the FishGenome Consortium. We thank Jinliang Wang for helpful support in the use of COLONY and EMIBD9 software. We thank Marta Cousido and Santiago Cerviño from the Spanish Institute of Oceanography (Vigo, Spain), as well the ICES WGBIE, for providing stock data for both hake Atlantic stocks. Finally, we also thank the two anonymous reviewers and the editor for their useful comments on the earlier version of the manuscript.

Authors contribution

Natalia Petit-Marty (Conceptualization [lead], Data curation [lead], Formal Analysis [lead], Funding acquisition [equal], Investigation [lead], Methodology [lead], Project administration [equal], Software [equal], Validation [equal], Visualization [equal], Writing – original draft [lead], Writing – review & editing [lead]), Laura Casas (Data curation [equal], Investigation [equal], Methodology [equal], Resources [equal], Writing – review & editing [equal]), Sara Rocha (Formal Analysis [equal], Investigation [equal], Methodology [equal], Software [equal], Validation [equal], Visualization [equal], Writing – review & editing [equal]), María Saura (Formal Analysis [equal], Investigation [equal], Methodology [equal], Validation [equal], Writing – review & editing [equal]), Inés González-Castellano (Formal Analysis [equal], Investigation [equal], Validation [equal], Writing – review & editing [equal]), Sergio Ramirez-Amaro (Investigation [equal], Methodology [equal], Resources [equal], Writing – review & editing [equal]), Erik Eschbach (Investigation [equal], Resources [equal], Writing – review & editing [equal]), Beatriz Guijarro (Investigation [equal], Resources [equal], Writing – review & editing [equal]), Reinhold Hanel (Investigation [equal], Resources [equal], Writing – review & editing [equal]), Armando Caballero (Conceptualization [equal], Formal Analysis [lead], Funding acquisition [equal], Investigation [equal], Methodology [lead], Software [equal], Validation [equal], Visualization [equal], Writing – review & editing [lead]), Fran Saborido-Rey (Conceptualization [equal], Data curation [equal], Funding acquisition [equal], Investigation [equal], Project administration [equal], Resources [lead], Writing – review & editing [lead]).

Supplementary data

Supplementary data is available at *ICES Journal of Marine Science* online.

Conflict of interest: The authors declare no conflict of interest.

Funding

NPM was supported by MSCA-PF 101066785 (FISHADAPT https://doi.org/10.3030/101066785). This study was framed within the Marine Science Programme (ThinkInAzul) supported by the Ministerio de Ciencia e Innovación and Xunta de Galicia with funding from the European Union NextGenerationEU (PRTR-C17.I1) and European Maritime and Fisheries Fund. IGC was funded by the Spanish Ministry of Universities (Margarita Salas postdoctoral fellowship, ref. RSU.UDC.MS17). DREAMER-1 was funded by the Spanish National R + D project "Dynamics of European hake recruitment" (CTM2015-66676-C2-1-R). The North Sea IBTS-Q3 and MEDITS surveys are co-funded by the European Union through the European Maritime and Fisheries Fund (EMFF) within the national program of collection, management, and use of data in the fisheries sector and support for scientific advice regarding the Common Fisheries Policy. This manuscript is built on scientific data and information collected under Service Contract "Improving cost-efficiency of fisheries research surveys and fish stocks assessments using next-generation genetic sequencing methods [EMFF/2018/015]" (EU Contract—EASME/EMFF/2017/1.3.2.10/SI2.790889) signed with the European Climate, Infrastructure and Environment Executive Agency (CINEA) and funded by the European Union.

Data availability

The data underlying this article are available in Zenodo, at https://zenodo.org/records/17839042. These datasets were derived from source read sequences stored at NCBI (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1329964/) under Bioproject PRJNA1329964.

References

Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*, 2009;**19**:1655–64. https://doi.org/10.1101/gr.094052.109

Allendorf FW, England PR, Luikart G *et al.* Genetic effects of harvest on wild animal populations. *Trends Ecol Evol* 2008;**23**:327–37. https://doi.org/10.1016/j.tree.2008.02.008

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

Árnason E, Koskela J, Halldórsdóttir K *et al.* Sweepstakes reproductive success via pervasive and recurrent selective sweeps. *eLife* 2023;**12**:e80781. https://doi.org/10.7554/eLife.80781

Caballero A. *Quantitative Genetics*. Cambridge University Press. 2020. 336. https://doi.org/10.1017/9781108630542

Caballero A. Developments in the prediction of effective population size. *Heredity*, 1994;**73**:657–79. https://doi.org/10.1038/hdy.1994.174

Catchen JM, Amores A, Hohenlohe P *et al.* Stacks: building and genotyping loci de novo from short-read sequences. *G3 (Bethesda)* 2011;**1**:171–82. https://doi.org/10.1534/g3.111.000240

Ceballos G, Ehrlich PR, Dirzo R. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proc Natl Acad Sci*, 2017;**114**:E6089–96. https://doi.org/10.1073/pnas.1704949114

Chang CC, Chow CC, Tellier LC *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 2015;**4**:s13742–015. https://doi.org/10.1186/s13742-015-0047-8

Clarke SH, Lawrence ER, Matte JM *et al.* Global assessment of effective population sizes: consistent taxonomic differences in meeting the 50/500 rule. *Mol Ecol* 2024;**33**:e17353. https://doi.org/10.1111/mec.17353

Danecek P, Bonfield JK, Liddle J *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* 2021;**10**:giab008. https://doi.org/10.1093/gigascience/giab008

Etter PD, Bassham S, Hohenlohe PA *et al.* SNP discovery and genotyping for evolutionary genetics using RAD sequencing. In: *Methods Mol Biol* 2011;**772**:157–78. https://doi.org/10.1007/978-1-61779-228-1_9

Excoffier L, Marchi N, Marques DA *et al.* fastsimcoal2: demographic inference under complex evolutionary scenarios. *Bioinformatics* 2021;**37**:4882–5. https://doi.org/10.1093/bioinformatics/btab468

FAO. *The State of World Fisheries and Aquaculture 2022: Towards Blue Transformation*. Rome: FAO. 2022. https://openknowledge.fao.org/items/11a4abd8-4e09-4bef-9c12-900fb4605a02. Accessed September 2024

Fedorca A, Mergeay J, Akinyele AO *et al.* Dealing with the complexity of effective population size in conservation practice. *Evolutionary Applications* 2024;**17**:e70031. https://doi.org/10.1111/eva.70031

Fernández-Míguez M, Pita A, Gomez A *et al.* Temporal uncoupling between demographic and genetic metrics in fisheries assessment: the European hake case study. *Frontiers in Marine Science* 2023;**10**:1214469. https://doi.org/10.3389/fmars.2023.1214469

Fogarty MJ, Botsford LW. Population connectivity and spatial management of marine fisheries. *Oceanography*, 2007;**20**:112–23. https://doi.org/10.5670/oceanog.2007.34

Frankham R. Effective population size/adult population size ratios in wildlife: a review. *Genetics Research*, 2007;**89**:491–503. https://doi.org/10.1017/S0016672308009695

Fumagalli M, Vieira FG, Korneliussen TS *et al.* Quantifying population genetic differentiation from next-generation sequencing data. *Genetics* 2013;**195**:979–92. https://doi.org/10.1534/genetics.113.154740

Gandra M, Assis J, Martins MR *et al.* Reduced global genetic differentiation of exploited marine fish species. *Mol Biol Evol* 2021;**38**:1402–12. https://doi.org/10.1093/molbev/msaa299

Garcia-Erill G, Albrechtsen A. Evaluation of model fit of inferred admixture proportions. *Mol Ecol Resour*, 2020;**20**:936–49. https://doi.org/10.1111/1755-0998.13171

Gargiulo R, Decroocq V, González-Martínez SC *et al.* Estimation of contemporary effective population size in plant populations: limitations of genomic datasets. *Evolutionary Applications* 2024;**17**:e13691. https://doi.org/10.1111/eva.13691

Hauser L, Carvalho GR. Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. *Fish and Fisheries*, 2008;**9**:333–62. https://doi.org/10.1111/j.1467-2979.2008.00299.x

Hidalgo M, Olsen EM, Ohlberger J *et al.* Contrasting evolutionary demography induced by fishing: the role of adaptive phenotypic plasticity. *Ecol Appl* 2014;**24**:1101–14. https://doi.org/10.1890/12-1777.1

Hixon MA, Johnson DW, Sogard M. BOFFFFs: on the importance of conserving old-growth age structure in fishery populations. *ICES J Mar Sci*, 2014;**71**:2171–85. https://doi.org/10.1093/icesjms/fst200

Hoban S, Bruford M, Jackson JDU *et al.* Genetic diversity targets and indicators in the CBD post-2020 global biodiversity framework must be improved. *Biol Conserv* 2020;**248**:108654. https://doi.org/10.1016/j.biocon.2020.108654

Hoey JA, Able KW, Pinsky ML. Genetic decline and recovery of a demographically rebuilt fishery species. *Mol Ecol*, 2022;**31**:5684–98. https://doi.org/10.1111/mec.16697

Hudson RR. Generating samples under a Wright-Fisher neutral model. *Bioinformatics*, 2002;**18**:337–8. https://doi.org/10.1093/bioinformatics/18.2.337

ICES. *ICES Advice 2024*. https://ices-library.figshare.com/collections/: ICES Advice Publications. 2024. https://doi.org/10.17895/ices.pub.c.6976944

Jones OR, Wang J. COLONY: a program for parentage and sibship inference from multilocus genotype data. *Mol Ecol Resour*, 2010;**10**:551–5. https://doi.org/10.1111/j.1755-0998.2009.02787.x

Korneliussen TS, Albrechtsen A, Nielsen R. ANGSD: analysis of next generation sequencing data. *BMC Bioinf*, 2014;**15**:1–13. https://doi.org/10.1186/s12859-014-0356-4

Kozlov A, Alves JM, Stamatakis A *et al.* CellPhy: accurate and fast probabilistic inference of single-cell phylogenies from scDNA-seq data. *Genome Biol* 2022;**23**:37. https://doi.org/10.1186/s13059-021-02583-w

Kuparinen A, Hutchings JA, Waples RS. Harvest-induced evolution and effective population size. *Evolutionary Applications*, 2016;**9**:658–72. https://doi.org/10.1111/eva.12373

Leone A, Álvarez P, García D *et al.* Genome-wide SNP based population structure in European hake reveals the need for harmonizing biological and management units. *ICES J Mar Sci* 2019;**76**:2260–6. https://doi.org/10.1093/icesjms/fsz161

Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. 2013. arXiv preprint. (arXiv:1303.3997)

Liu X, Fu YX. Stairway Plot 2: demographic history inference with folded SNP frequency spectra. *Genome Biol*, 2020;**21**:280. https://doi.org/10.1186/s13059-020-02196-9

Lowerre-Barbieri S, DeCelles G, Pepin P, *et al.* Reproductive resilience: A paradigm shift in understanding spawner-recruit systems in exploited marine fish. *Fish and Fisheries* 2017;**18**:285–312.

Lynch M, Conery JS. The origins of genome complexity. *Science*, 2003;**302**:1401–4. https://doi.org/10.1126/science.1089370

Manel S, Guerin PE, Mouillot D *et al.* Global determinants of freshwater and marine fish genetic diversity. *Nat Commun* 2020;**11**:1–9. https://doi.org/10.1038/s41467-020-14409-7

Marandel F, Charrier G, Lamy JB *et al.* Estimating effective population size using RADseq: effects of SNP selection and sample size. *Ecology and Evolution* 2020;**10**:1929–37. https://doi.org/10.1002/ece3.6016

Marandel F, Lorance P, Berthelé O *et al.* Estimating effective population size of large marine populations, is it feasible? *Fish and Fisheries* 2019;**20**:189–98. https://doi.org/10.1111/faf.12338

Martinez P, Casas L, Petit-Marty N *et al.* Screening a new European hake (Merluccius merluccius) chromosome-level genome assembly suggests an XX/XY sex-determining system driven by the SRY-box transcription factor 3 (sox3). G3: genes, Genomes. *G3: Genes, Genomes, Genetics* 2025;**15**:jkaf127. https://doi.org/10.1093/g3journal/jkaf127

Milano I, Babbucci M, Cariani A *et al.* Outlier SNP markers reveal fine-scale genetic structuring across European hake populations (Merluccius merluccius). *Mol Ecol* 2014;**23**:118–35. https://doi.org/10.1111/mec.12568

Nei M. *Molecular Evolutionary Genetics*. New York: Columbia University Press. 1987https://doi.org/10.7312/nei-92038

Nunziata SO, Weisrock DW. Estimation of contemporary effective population size and population declines using RAD sequence data. *Heredity*, 2018;**120**:196–207. https://doi.org/10.1038/s41437-017-0037-y

Ovenden JR, Leigh GM, Blower DC *et al.* Can estimates of genetic effective population size contribute to fisheries stock assessments? *J Fish Biol* 2016;**89**:2505–18. https://doi.org/10.1111/jfb.13129

Palstra FP, Fraser DJ. Effective/census population size ratio estimation: a compendium and appraisal. *Ecology and Evolution*, 2012;**2**:2357–65. https://doi.org/10.1002/ece3.329

Petit-Marty N, Liu M, Tan IZ *et al.* Declining population sizes and loss of genetic diversity in commercial fishes: a simple method for a first diagnostic. *Frontiers in Marine Science* 2022;**9**:872537. https://doi.org/10.3389/fmars.2022.872537

Piñeiro C, Saínza M. Age estimation, growth and maturity of the European hake (Merluccius merluccius (Linnaeus, 1758)) from Iberian Atlantic waters. *ICES J Mar Sci*, 2003;**60**:1086–102. https://doi.org/10.1016/S1054-3139(03)00086-9

Pinsky ML, Eikeset AM, Helmerson C *et al*. Genomic stability through time despite decades of exploitation in cod on both sides of the Atlantic. *Proc Natl Acad Sci* 2021;**118**:e2025453118. https://doi.org/10.1073/pnas.2025453118

Pinsky ML, Palumbi SR. Meta-analysis reveals lower genetic diversity in overfished populations. *Mol Ecol*, 2014;**23**:29–39. https://doi.org/10.1111/mec.12509

Pita A, Pérez M, Balado M *et al*. Out of the Celtic cradle: the genetic signature of European hake connectivity in south-western Europe. *J Sea Res* 2014;**93**:90–100. https://doi.org/10.1016/j.seares.2013.11.003

Pita A, Pérez M, Velasco F *et al*. Trends of the genetic effective population size in the southern stock of the European hake. *Fish Res* 2017;**191**:108–19. https://doi.org/10.1016/j.fishres.2017.02.022

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. 2024. https://www.r-project.org. Accessed November 2024.

Rambaut A. *Figtree*, version 1.4.4. Institute of Evolutionary Biology, University of Edinburgh. 2018. https://tree.bio.ed.ac.uk/software/figtree. Accessed September 2024.

Saborido-Rey F, Junquera S. Spawning biomass variation in Atlantic cod (Gadus morhua) in Flemish Cap in relation to changes in growth and maturation. *Journal of Northwest Atlantic Fishery Science* 1999;**25**:83–90. https://doi.org/10.2960/J.v25.a7

Santiago E, Caballero A, Köpke C *et al*. Estimation of the contemporary effective population size from SNP data while accounting for mating structure. *Mol Ecol Resour* 2024;**24**:e13890. https://doi.org/10.1111/1755-0998.13890

Santiago E, Köpke C, Caballero A. Accounting for population structure and data quality in demographic inference with linkage disequilibrium methods. *Nat Commun* 2025;**16**:6054. https://doi.org/10.1038/s41467-025-61378-w

Santiago E, Novo I, Pardiñas AF *et al*. Recent demographic history inferred by high-resolution analysis of linkage disequilibrium. *Mol Biol Evol* 2020;**37**:3642–53. https://doi.org/10.1093/molbev/msaa169

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

Spielman D, Brook BW, Frankham R. Most species are not driven to extinction before genetic factors impact them. *Proc Natl Acad Sci* 2004;**101**:15261–4. https://doi.org/10.1073/pnas.0403809101

Staby A, Skjæraasen JE, Geffen AJ *et al*. Spatial and temporal dynamics of European hake (Merluccius merluccius) in the North Sea. *ICES J Mar Sci* 2018;**75**:2033–44. https://doi.org/10.1093/icesjms/fsy103

STECF. European Commission, Joint Research Centre, Scientific, Technical and Economic Committee for Fisheries. 2025. *Methodologies for Mediterranean Stock Assessments. (STECF-25-01)*. https://publications.jrc.ec.europa.eu/repository/handle/JRC142274

Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 1989;**123**:585–95. https://doi.org/10.1093/genetics/123.3.585

Venu V, Harjunmaa E, Dreau A *et al*. Fine-scale contemporary recombination variation and its fitness consequences in adaptively diverging stickleback fish. *Nat Ecol Evol* 2024;**8**:1337–52. https://doi.org/10.1038/s41559-024-02434-4

Wang J, Santiago E, Caballero A. Prediction and estimation of effective population size. *Heredity*, 2016;**117**:193–206. https://doi.org/10.1038/hdy.2016.43

Wang J. EMIBD9: estimating 9 condensed IBD coefficients, inbreeding and relatedness from marker genotypes. *Heredity* 2025;**134**:155–61. https://doi.org/10.1038/s41437-024-00739-5

Waples RS. Practical application of the linkage disequilibrium method for estimating contemporary effective population size: a review. *Mol Ecol Resour* 2024;**24**:e13879. https://doi.org/10.1111/1755-0998.13879

Waples RS. The idiot's guide to effective population size. *Mol Ecol*, 2025;**34**:e17670. https://doi.org/10.1111/mec.17670

Watterson GA. On the number of segregating sites in genetical models without recombination. *Theor Popul Biol* 1975;**7**:256–76. https://doi.org/10.1016/0040-5809(75)90020-9

Westgaard JI, Staby A, Aanestad Godiksen J *et al*. Large and fine scale population structure in European hake (Merluccius merluccius) in the Northeast Atlantic. *ICES J Mar Sci* 2017;**74**:1300–10. https://doi.org/10.1093/icesjms/fsw249

Wright S. Evolution in mendelian populations. *Genetics* 1931;**16**:97–159. https://doi.org/10.1093/genetics/16.2.97

Wright S. The genetical structure of populations. *Annals of Eugenics* 1949;**15**:323–54. https://doi.org/10.1111/j.1469-1809.1949.tb02451.x

Handling editor: John Horne