

## RESEARCH ARTICLE

Diversity and Distributions WILEY

# Ecological niche modelling and population genomics provide insights into the geographic and demographic 'explosion' of a non-indigenous salmonid

Simo Njabulo Maduna  | Paul Eric Aspholm  | Ane-Sofie Bednarczyk Hansen  |  
Cornelya F. C. Klütsch  | Snorre B. Hagen 

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

## Correspondence

Simo Njabulo Maduna, Department of Ecosystems in the Barents Region, Svanhovd Research Station, Norwegian Institute of Bioeconomy Research, Svanvik 9925, Norway.

Email: [simo.maduna@nibio.no](mailto:simo.maduna@nibio.no)

## Funding information

Klima- og miljødepartementet; Norsk institutt for Bioøkonomi

Editor: Zhonghe Hou

## Abstract

**Aim:** Effective management of non-indigenous species requires knowledge of their dispersal factors and founder events. We aim to identify the main environmental drivers favouring dispersal events along the invasion gradient and to characterize the spatial patterns of genetic diversity in feral populations of the non-native pink salmon within its epicentre of invasion in Norway.

**Location:** Mainland Norway and North Atlantic Basin.

**Methods:** We first conducted SDM using four modelling techniques with varying levels of complexity, which encompassed both regression-based and tree-based machine-learning algorithms, using climatic data from the present to 2050. Then, we used the triple-enzyme restriction-site associated DNA sequencing (3RADseq) approach to genotype over 30,000 high-quality single-nucleotide polymorphisms to elucidate the patterns of genetic diversity and gene flow within the pink salmon putative invasion hotspot.

**Results:** We discovered temperature- and precipitation-related variables drove pink salmon distributional shifts across its non-native ranges and that climate-induced favourable areas will remain stable for the next 30 years. In addition, all SDMs identified north-eastern Norway as the epicentre of the pink salmon invasion, and genomic data revealed that there was minimal variation in genetic diversity across the sampled populations at a genome-wide level in this region. While utilizing a specific group of 'diagnostic' SNPs, we observed a significant degree of genetic differentiation, ranging from moderate to substantial, and detected four hierarchical genetic clusters concordant with geography.

**Main Conclusions:** Our findings suggest that fluctuations in climate extreme events associated with ongoing climate change will likely maintain environmental favourability for the pink salmon outside its 'native'/introduced ranges. Locally invaded rivers are themselves potential source populations of invaders in the ongoing secondary spread of pink salmon in Northern Norway. Our study shows

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *Diversity and Distributions* published by John Wiley & Sons Ltd.

that SDMs and genomic data can reveal species distribution determinants and provide indicators to aid in post-control measures and potentially inferences about their success.

#### KEYWORDS

adaptive management, invasive bridgehead effect, non-indigenous species, *Oncorhynchus gorbuscha*, pink salmon, species distribution modelling

## 1 | INTRODUCTION

An ecological explosion means the enormous increase in numbers of some kind of living organism—it may be[...]a green plant like the prickly pear, or an animal like the grey squirrel. I use the word “explosion” deliberately, because it means the bursting out from control forces that were previously held in restraint by other forces.

Elton (1958: 15)

Biological invasions are a major challenge of global environmental change, where the introduction and establishment of non-indigenous species (NIS) beyond their natural distributional range is occurring at unprecedented rates around the world (Bax et al., 2003; Shackleton et al., 2019; Simberloff et al., 2013). The impact posed by NIS on the communities and ecosystems they invade is substantial and complex (Bellard et al., 2016; Giakoumi et al., 2019; Shackleton et al., 2019; Simberloff et al., 2013), which vary in degree (negative, neutral, positive), magnitude (benign, moderate, severe) and scale (individual, community, ecosystem) (Mrnak et al., 2023). While only a small proportion of introduced NIS become invasive in the recipient habitat, community or ecosystem (Blackburn et al., 2011), invaders can exert severe per-capita effects on local communities, causing detrimental ecological and economic impacts (Giakoumi et al., 2019; Shackleton et al., 2019). Invasive NIS can have an immediate impact on recipient communities through direct ecological interactions, including competition, predation and disease transmission, as well as an indirect impact by modifying habitats (e.g. reduced plant cover, increased soil erosion, increased eutrophication, pollutant transference), community structure and dynamics and possibly disrupting their suitability (Dunlop et al., 2021; Lurgi et al., 2016; Simberloff et al., 2005). Furthermore, invasive NIS can have large impacts on the genetic integrity of indigenous species through hybridization, introgression and selection, causing further shifts in the evolutionary pathway of native species and posing a threat to species integrity (Allendorf & Lundquist, 2003; Echelle & Echelle, 1997; Thirstrup et al., 2015). As a result, actions that prevent species establishments before they occur as well as post-establishment control measures have become major conservation priorities, with significant investments in invasive species and biosecurity research in the recent years (Mellin et al., 2016; Sandlund et al., 2019). Effective and successful management of NIS requires an understanding of how they were introduced, how they spread and how they are genetically structured.

Ecological niche and species distribution models (SDMs) provide an analytical framework for interpolating and extrapolating species distributions by linking, respectively, physiological (mechanistic models) or distribution data (correlative models) to environmental predictor variables that contribute to a species' survival and propagation (Araújo et al., 2019; Sillero & Barbosa, 2021; Václavík & Meentemeyer, 2009). The objective of SDMs in the context of NIS is to identify the environmental factors that promote or inhibit the establishment of NIS in non-native ranges and to predict their distribution and suitable habitats, as well as to determine their potential for spread factoring in the mode and strength of dispersal capacity (Burgess et al., 2022; El-Barougy et al., 2021; Melo-Merino et al., 2020; Rödder et al., 2009). Moreover, already built SDMs can be projected as well to other scenarios in space and time (Barbosa et al., 2009; Sillero & Barbosa, 2021; Werkowska et al., 2017; Yates et al., 2018) to further uncover broader suitable habitats in the face of ongoing anthropogenic pressures (Burgess et al., 2021, 2022; El-Barougy et al., 2021; Melo-Merino et al., 2020; Rödder et al., 2009). Such data are essential for making informed management decisions, such as identifying and prioritizing infestation-prone habitats (Melo-Merino et al., 2020) and planning post-establishment interventions such as harvesting, culling (Altenritter et al., 2022; Sandlund et al., 2019) or disrupting dispersal routes through restricting access (e.g. fencing, trapping) in conjunction with harvesting and/or culling (Burgess et al., 2022; Letnic et al., 2015). Therefore, SDMs have become popular tools for nature conservation and management, and their robustness can be increased by incorporating population genetics and genomics data to account for the adaptive capacity of populations distributed across a heterogeneous landscape (Burgess et al., 2021; Bush et al., 2016; Mathieu-Bégné et al., 2021; Scoble & Lowe, 2010).

Population genetics and genomics have been central for understanding biological invasions by providing frameworks for unravelling concealed patterns and processes involved with NIS introductions and invasions (Burgess et al., 2021; Estoup et al., 2016; Matheson & McGaughan, 2022), and advancing contentious scientific debates (e.g. genetic paradox in invasion biology—the success of invasive species despite bottlenecks: Buchholz et al., 2023; Colautti et al., 2017; Dlugosch et al., 2015; Estoup et al., 2016; Frankham, 2005; Roman & Darling, 2007; Sakai et al., 2001). Applied population genomics can reveal the demographic history of NIS, which can identify and/or validate its founders or putative introduction sources. In combination with historical, ecological and indigenous knowledge, genomic tools can be very informative

and influence management decisions by highlighting important information regarding population dynamics and biological invasion processes (Burgess et al., 2021; Hogg et al., 2022; Matheson & McGaughan, 2022). Invading or introduced species often have low genetic diversity in their pioneer populations because of founder effects (Colautti et al., 2017; Dlugosch et al., 2015; Estoup et al., 2016). Since normally there is no hybridization in this first wave of invaders, genetic diversity among these individuals is projected to be much lower than in the founder population. Individuals can be assigned to specific populations based on their genetic make-up, and major disruptions to gene flow can be identified, which is useful information in the fight against subsequent dispersal events along invasion (dispersal) gradients or 'invasion front'.

Pink salmon *Oncorhynchus gorbuscha* (Walbaum, 1792) is one of six Pacific salmon species in the family Salmonidae (Radchenko et al., 2018; Scott & Crossman, 1973). Pink salmon have a 2-year lifespan, during which they spend about 1 year at sea before returning to native rivers to spawn at age 2 in odd or even years (Quinn, 2005). The temporal isolation (allochronic) between odd-year and even-year spawners has led to the evolution of reproductively isolated year-classes, or -lineages (Aspinwall, 1974; Christensen et al., 2021; Quinn, 2005). Genome-wide studies showed that genetic divergence is greater between broodlines rather than geographically (Christensen et al., 2021; Limborg et al., 2014; Tarpey et al., 2018) but not exceeding the general level of inter-population differences in salmonids (Althukov et al., 2000; Sandlund et al., 2019). Pink salmon is an economically valuable, heavily exploited species in its native range, so conservation efforts have been made to maintain its exploitation levels. Pink salmon is a native of the southern Arctic Ocean (between North America and Asia) and the northern Pacific Ocean (between Alaska and British Columbia, on Russian side south to Japan), but it is increasingly expanding outside its native and introduced ranges due to translocation and rising temperature (Brenner et al., 2012; Hendry et al., 2004; Quinn, 2005; Sandlund et al., 2019; Ueda, 2012).

In the late 1950s, pink salmon was transferred from the northern Pacific Ocean and introduced it to the Laurentian Great Lakes and White Sea drainage basins of North-West Russia. Pink salmon has been detected in Norwegian waters since the late 1950s, after the first successful introduction of pink salmon fry in rivers draining to the White Sea. Pink salmon in Norway has a predominantly odd-year broodline, with even-year fish being far less abundant. During the last 5–10 years, the odd-year broodline of the 2-year life salmonid has shown a rapid and possibly climate-driven geographic and demographic expansion, with an unanticipated peak in 2017 in Norway and other parts of the North Atlantic area (Armstrong et al., 2018; Millane et al., 2019; Mo et al., 2018; Sandlund et al., 2019). In Norway, in the year 2017, local initiatives were implemented with the aim of capturing and eradicating the pink salmon (Mo et al., 2018). This action was taken in response to the official designation of the species as an undesirable invasive species, despite the limited scientific evidence available at the time (Jonsson & Jonsson, 2018). However, later in 2021, the pink salmon was observed in rivers from north to

south, along the entire coastline of Norway (Berntsen et al., 2022; Diaz Pauli et al., 2023). Secondary spread and establishment to locations other than the original target release area(s) is now an ongoing major challenge for aquatic and natural resource management in Norway and other northern European countries because it becomes impossible to contain unwanted species with a high rate of long-distance straying, such as pink salmon (Brenner et al., 2012; Hendry et al., 2004; Quinn, 2005; Ueda, 2012).

Here, we used SDM and population genomics to investigate mechanisms that shape the distribution of the range-expanding pink salmon in northern Europe. We first conducted SDM using four modelling techniques with varying levels of complexity, which encompassed both regression-based and tree-based machine-learning algorithms, using climatic data from the present to 2050. Then, we used the triple-enzyme restriction site-associated DNA sequencing (3RADseq) approach of Bayona-Vásquez et al. (2019) to genotype high-quality single-nucleotide polymorphisms (SNPs) in several sampling populations collected from different river systems within the invasion hotspot of the pink salmon in Arctic Norway. We specifically aimed to (1) identify the key climatic drivers of pink salmon distribution during its reproductive stage, (2) determine the climatic favourable habitats within and beyond the current non-native and introduced ranges in northern Europe, (3) determine the extent of projected distributional ranges under a climate scenario by the 2050s, (4) understand how genetic drift and gene flow influenced the spatial genetic patterns of the pink salmon following the 2017 surge and culling events within the invasion hotspot. We anticipate that such baseline data will aid in making genomic-informed management decisions on the species' spatial invasion gradient from north-east (Russia) to north-west (Norway).

## 2 | METHODS

### 2.1 | Species distribution modelling

We inferred climatically favourable areas for the pink salmon under current and future climate with ecological niche and species distribution models (SDMs) based on this species' current distribution in Norwegian rivers on presence data obtained during its spawning period. In terms of criteria for handling and processing of response and predictor variables, as well as model creation and evaluation, we followed best-practice standards in ecological niche and species distribution modelling (Araújo et al., 2019; Sillero & Barbosa, 2021). As Appendix S2, we present a description of the six modelling stages used to construct our SDMs in accordance with the ODMAP (Overview, Data, Model, Assessment and Prediction) protocol of Zurell et al. (2020).

Although prior research has emphasized the significance of utilizing both the native and invaded ranges for predicting the spread of invasive species (e.g. Zhang et al., 2020 and references therein), we opted to focus on the invaded range restricted to Norwegian estuarine and freshwater ecosystems. This is because the pink salmon

in Atlantic waters were introduced over six decades ago (~30 generations) and in eastern part of North (or North-eastern) Norway, they constitute already established, locally adapted feral populations (Bjerknes & Waag, 1980; Gordeeva et al., 2015; Niemelä et al., 2016; Sandlund et al., 2019; Zubchenko et al., 2004), which are likely themselves the source of additional new introductions rather than from invaders originating from the native introduced range. This self-accelerating process whereby invasion begets invasion, that is, the so-called invasive bridgehead effect, is not that uncommon for highly invasive species (Bertelsmeier & Keller, 2018; Lombaert et al., 2010). Consequently, while global occurrence data may provide for a more accurate assessment of the whole potential invasive range, it is more important from a management standpoint to predict the places most likely to be invaded next (Baquero et al., 2021; Barbet-Massin et al., 2018). Indeed, SDMs built on NIS occurrences only at invaded areas have proven capable of producing accurate forecasts and providing actionable information for the management of invasive species (Baquero et al., 2021; Barbet-Massin et al., 2018; Muñoz & Real, 2006; Pereira et al., 2020). Occurrence and environmental data retrieval and processing and distribution modelling were performed with R v.4.1.3 (R Core Team, 2022) using custom R scripts.

### 2.1.1 | Occurrence data

We collected occurrence data of pink salmon in Norway from the GBIF—the Global Biodiversity Information Facility—biodiversity database (<https://www.gbif.org/>). These records were complemented with additional information from invasive species monitoring reports (Berntsen et al., 2018, 2020, 2022; Husa et al., 2022) and peer-reviewed scientific literature (Diaz Pauli et al., 2023; Sandlund et al., 2019). We mapped the cleaned data and visually inspected points of occurrences to identify and remove occurrence data with coordinate uncertainty >10 km, given that our target spatial resolution of grid cells was 10 km × 10 km (i.e. 100 km<sup>2</sup>). Additionally, we cleaned species occurrence data by removing duplicates, records of 'absence' (with abundance or individual count reported as zero) and occurrences with missing and/or unlikely coordinates. The final database included 300 UTM cell occurrences, where each occupied cell counted as only one presence, regardless of the overall number of observations per cell. This is equivalent to thinning occurrence data at 10-km intervals, and it lowers the impact of sample bias caused by disproportionately extensive survey efforts in few particularly accessible places (Baquero et al., 2021; Pereira et al., 2020). Given the invasion history of pink salmon in Norway, we considered any remaining spatial autocorrelation in the data, that is, 100-km<sup>2</sup> cells having a higher chance of presence because they are close to other 100-km<sup>2</sup> cells with presence, is a natural effect of dispersal processes. Consequently, we did not intentionally remove such cells for downstream distribution modelling (Baquero et al., 2021; Legendre & Legendre, 2012).

### 2.1.2 | Climatic variables

We chose 19 'bioclimate' layers and one topographic variable (elevation) as predictors for building SDMs, which depict the most comprehensive environmental drivers for the selection of rivers for spawning and nursery areas (Table 1). We used the R package *SDMPREDICTORS* v.0.2.13 (Bosch & Fernandez, 2022) to access and download environmental data from WorldClim (<http://www.worldclim.org/>; Fick & Hijmans, 2017; Hijmans et al., 2005). To mitigate multicollinearity, we employed the *corSelect()* function from the R program *FUZZYSIM* v.4.3 (Barbosa, 2015). We set a threshold value for the correlation coefficient (*cor.thresh*) at >0.7 and eliminated the variable with the highest Variance Inflation Factor (VIF: Mansfield & Helms, 1982; Marquardt, 1970) among the correlated variables. Similarly, to assess the probable future distribution in pink salmon invasive range, we utilized the climate projections from WorldClim. All environmental variables except elevation were obtained from the representative concentration pathway (RCP) 2.6 scenario (minimum emission hypothesis) under the global climate model community earth system model (CCSM4) (Gent et al., 2011) for the period 2050 (average for 2041–2060).

### 2.1.3 | Model building

We fitted and validated models using an ensemble approach with four SDM methods of varying complexity, two regression-based (i–ii) and two machine-learning approaches (iii–iv): (i) generalized linear models (GLM: McCullagh & Nelder, 1989), (ii) generalized additive models (GAM: Hastie & Tibshirani, 1990) with smoothing splines, (iii) boosted regression trees (BRT: Elith et al., 2008, 2010) and (iv) maximum entropy (*MaxEnt*: Elith et al., 2010; Phillips et al., 2006). We fitted the models using the R package *BASE* v.4.1.3 (R Core Team, 2022), *GAM* v.1.20.2 (Hastie, 2022), *GBM* v.2.1.8.1 (Greenwell et al., 2022) and *MAXNET* v.0.1.4 (Phillips, 2021), respectively. For the machine learning and predictive modelling technique, BRT, we optimized the learning rate (*shrinkage*) and tree complexity (*interaction.depth*), while keeping the bag fraction (*bag.fraction*=0.5) and number of trees (*n.trees*=1000) constant to fit a good model. The learning rate is used to govern how much each individual tree contributes to the larger model. The number of nodes in a tree is adjusted by tree complexity, which governs the model's interaction complexity (e.g. if tree complexity is 3, up to three-way interactions can be fit). During preliminary model runs, we evaluated differing combinations of learning rate and tree complexity values, and determined that *shrinkage*=0.1 and *interaction.depth*=2 were the parsimonious optimum parameters for building the BRT model.

### 2.1.4 | Model validation

To evaluate model performance, we initially partitioned the data into fivefold cross-validation sets: the grid cells of the study area were

TABLE 1 Variables used for modelling climatic favourability modelling of pink salmon.

Code	Definition	Units
BIO01	Annual mean temperature	Degrees celsius (°C)
BIO02 <sup>a</sup>	Mean diurnal range [Mean of monthly (max temp – min temp)]	°C
BIO03 <sup>a</sup>	Isothermality (BIO2/BIO7 * 100)	Percentage (%)
BIO04	Temperature seasonality (standard deviation * 100)	°C
BIO05	Max temperature of warmest month	°C
BIO06	Min temperature of coldest month	°C
BIO07	Temperature annual range (BIO5–BIO6)	°C
BIO08 <sup>a</sup>	Mean temperature of wettest quarter	°C
BIO09	Mean temperature of driest quarter	°C
BIO10	Mean temperature of warmest quarter	°C
BIO11	Mean temperature of coldest quarter	°C
BIO12	Annual precipitation	Millimetres (mm)
BIO13	Precipitation of wettest month	mm
BIO14	Precipitation of driest month	mm
BIO15 <sup>a</sup>	Precipitation seasonality (coefficient of variation)	Fraction
BIO16	Precipitation of wettest quarter	mm
BIO17	Precipitation of driest quarter	mm
BIO18 <sup>a</sup>	Precipitation of warmest quarter	mm
BIO19	Precipitation of coldest quarter	mm
Elevation <sup>a</sup>	Topography—elevation	Metres (m)

<sup>a</sup>Final variables used for modelling the climatically favourable habitats.

split into five groups, or *folds*, using R package *BLOCKCV* v.2.1.4 (Valavi et al., 2018), with a systematic selection and assignment of spatial blocks into 10 100×100km<sup>2</sup> folds. The tactic of spatially separating our data into training and test data sets enabled us to determine if models perform effectively in both nearby and distant areas, which is especially crucial if models are to be extrapolated in space or time (Baquero et al., 2021; Valavi et al., 2018). This resulted in five different SDMs in which each fold was sequentially left out of the training data and used for validation of model predictions (Fielding & Bell, 1997), where an overall total of 20 models (4 modelling algorithms × 5 replicates) were computed in the cross-validation phase. We then used the R package *MODeVA* v.3.5 (Barbosa et al., 2013) to assess the cross-validation performance of each modelling algorithm. Following Lake et al. (2020)'s proposal to employ various and diverse metrics for model evaluation, we selected three metrics that focus on different aspects of model performance: (i) the Area Under the receiver operating characteristic Curve (AUC) that assesses overall discrimination performance, that is, the model's ability to distinguish between presence and absence localities by assigning them higher predicted values; (ii) the True Skill Statistic (TSS), which assesses classification performance by calculating the proportion of correctly classified presences and absences (Allouche et al., 2006), with training prevalence as the threshold value for identifying predicted presences and predicted absences; and (iii) Miller's calibration slope that evaluates model reliability, defined as the total deviation of predicted probabilities from observed occurrence frequencies

(Miller et al., 1991; Pearce & Ferrier, 2000). The values of these metrics were then averaged across the five cross-validation folds, and models whose average cross-validation performance fell below an acceptable threshold for any of the model evaluation criteria were eliminated from further analysis (Baquero et al., 2021). This performance threshold was set to 0.7 for AUC, which limits weak predictions (Baquero et al., 2021; Rapacciuolo et al., 2012; Swets, 1988). Given that AUC can vary from 0 to 1 and TSS can range from –1 to 1, the TSS proportional performance threshold was set to 0.4. We selected a performance criterion of 0.5 above or below the ideal value of 1 for Miller's calibration slope (Baquero et al., 2021).

## 2.1.5 | Model post-processing and projection

After cross-validation, we projected retained models to a broader Area across the Arctic and North Atlantic Oceans, including the Baltic and Mediterranean Seas, and to a future climate scenario. Given the sensitivity of model projections to removal of data points (Araújo et al., 2019), models were fitted again using the whole data set (i.e. without excluding validation data) in this instance. For all regression-based approaches, we transformed projected probability values to prevalence-independent favourability, which is a function of presence probability and the presence-absence ratio of the species in the modelled sample (Real et al., 2006), using the *Fav()* function of the R package *FUZZYSIM*. This eliminates



the effect of species prevalence, making the predicted values consistent and directly comparable across SDMs, species and time periods (Acevedo & Real, 2012; Real et al., 2017). We further used FUZZYSIM to calculate the similarity of the predicted habitat suitability and favourability between the best-performing models (*MaxEnt*, *GLM*, *GAM* and *BRT*) and the *ensemble* predictions for current and future climate scenarios. We used two niche comparison metrics of Warren et al. (2008), namely Schoener's *D* index (Schoener, 1968) and Warrant's *I* index, which range from 0 (no similarity or overlap) to 1 (complete similarity or overlap) (Warren et al., 2008, 2010). Moreover, we examine the extrapolation potential of the SDMs across the North Atlantic and Arctic Oceans using the Multivariate Environmental Similarity Surface (MESS: Elith et al., 2010), with the *mess()* function of the *DISMO* v.1.3-9 R package (Hijmans et al., 2022).

## 2.2 | Population genomics

### 2.2.1 | Sampling and DNA extraction

Following the 2017 surge and culling events within the pink salmon's invasion hotspot, we launched a small citizen science initiative in 2019 in which local river guards, fishermen and anglers provided tissue or fin clip samples for the species in its putative invasion hotspot. The sampling area was comprised of three river basins (Munkelva, Neidenelva and Klokkelva) and two sections of a fjord arm, Jarfjord (Jarfjord mouth and Jarfjordbotn), on Varanger Fjord in Sør-Varanger Municipality in Troms and Finnmark county, Norway (Table 2, Figure 1). It is likely that samples collected in fjords are fish that had not migrated to their spawning rivers or were a secondary wave of invaders. Nevertheless, we consider fjords as (unknown) 'river' basins. We used the DNeasy Tissue Kit following the manufacturer's instructions (Qiagen) to extract total genomic DNA from each sample. Then, we quantified all DNA stocks with the Quantus™ Fluorometer using the QuantiFluor® ONE dsDNA System (Promega). Finally, we normalized each DNA sample to 10 ng/μL in nuclease and ion-free water and stored the working stocks at -21°C prior to 3RAD library preparation.

### 2.2.2 | 3RAD library preparation and sequencing

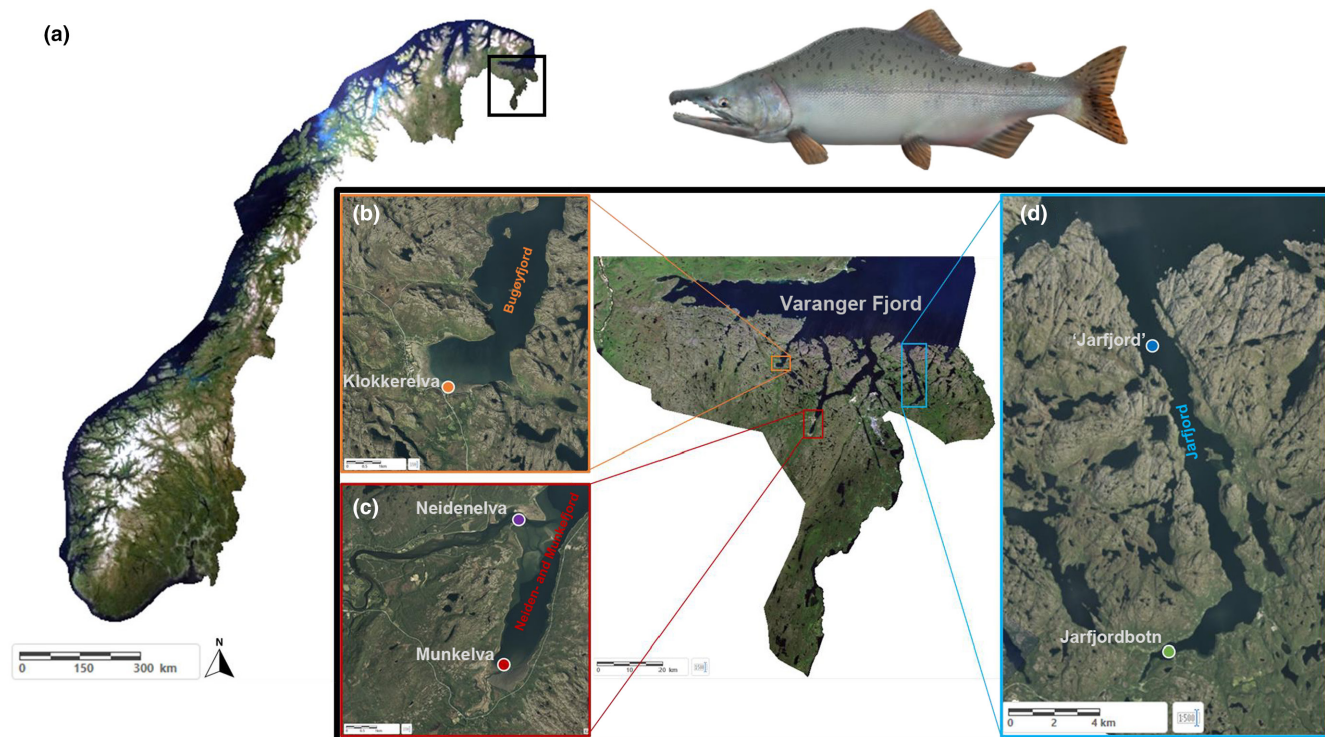
We prepared RADseq libraries using the Adapterama III library preparation protocol (Bayona-Vásquez et al., 2019; their Supplemental File S1), which is a modified version of double-digest (dd)RAD (Peterson et al., 2012) that uses three enzymes for digesting genomic DNA (3RAD). This procedure reduces the amount of DNA required since the third enzyme inhibits the formation of adapter dimers and DNA chimeras during the simultaneous digestion and ligation reactions (Bayona-Vásquez et al., 2019; Graham et al., 2015; Hoffberg et al., 2016). We surveyed the literature and identified the frequently used enzymes in RADseq experiments of fish, and we chose a pair of enzymes that were compatible with one of the 3RADseq designs (Bayona-Vásquez et al., 2019; their Table 1). We selected *MspI* (C|CGG) as the frequent cutter and *BamHI*-HF (G|GATCC) as the rare cutter, which fit the 3RAD Design 2 with *Clai* (AT|CGAT) as the third enzyme for suppressing phosphorylated ends in *MspI* recognition sites. A detailed account of the laboratory procedure of the 3RAD library preparation can be found within the Appendix S1. We included 10 technical replicates to estimate genotyping error rates. Cleaned and indexed library pools per design were sent to the Norwegian Sequencing Centre (NSC) for quality control and subsequent final size selection using a one-sided bead clean-up (0.7:1 ratio) to capture 550 bp ± 10% fragments, and the final paired-end (PE) 150 bp sequencing on one lane each of the Illumina HiSeq 4000 platform.

### 2.2.3 | Read mapping and variant calling

After assessing the quality of the sequencing run using FASTQC (Andrews, 2010) plus MULTIQC v.2.31 (Ewels et al., 2016), we processed the data as follows. First, we demultiplexed, cleaned and trimmed the raw sequences with the perl script *process\_radtags.pl* included as part of STACKS v.2.51 (Rochette et al., 2019). We ran the script using 'inline\_inline' mode, removed reads with an uncalled bases (-c) and discarded reads with low quality scores (-q) with a default sliding window of 15% of the length of the read and raw Phred score of

TABLE 2 Information about localities (geographic coordinates of the river mouth, catchment area and mean annual flow at the outlet of rivers; Sandlund et al., 2019) and sample sizes (*N*) for pink salmon from the respective Norwegian rivers and fjord.

Name	River mouth/inlet coordinates (degrees minutes, WGS84)	Catchment area (km <sup>2</sup> )	Annual mean flow (m <sup>3</sup> s <sup>-1</sup> )	<i>N</i>
(a) River (Varanger Fjord arm)				
Munkelva (Munkefjord)	69°38' N 29°27' E	410	2.0	14
Neidenelva (Neidenfjord)	69°42' N 29°24' E	2988	29.5	35
Klokkelva (Bugøyfjord)	69°51' N 29°23' E	142	1.2	13
(b) Fjord				
Upper Jarfjord	69°48.8' N 29°24.5' E	–	–	8
Jarfjordbotn	69°48.8' N 29°24.5' E	–	–	16
Total				86



**FIGURE 1** (a) Sampling map for the pink salmon *Oncorhynchus gorbusha* in Varanger Fjord, Norway, and image of a pink salmon obtained from Wiki for Fishing Planet ([https://wiki.fishingplanet.com/Pink\\_Salmon](https://wiki.fishingplanet.com/Pink_Salmon)). (b) The eastern-most Varanger Fjord arm Bugøyfjord, where sampling river network Klokkelva (orange point with white outline) drains. (c) The two adjacent Neiden- and Munkefjord arms of Varanger Fjord, where the central-eastern sampling river networks Neidenelva (purple point with white outline) and Munkelva (red point with white outline) discharge. (d) The western-most Varanger Fjord arm Jarfjord along which pink salmon was collected at two sampling sites within the upper and lower ends, 'Jarfjord' (blue point with white outline) and Jarfjordbotn (green point with white outline), respectively.

20. We rescued sequence tags (internal tags) and RAD-tags (enzyme over-hang) within two mismatches of their expected sequence (-r); otherwise, reads were discarded. We truncated (-t) PE150 reads to 140nt to have equal length among all reads with different barcodes. Second, we downloaded the chromosomal-level reference genome of the pink salmon (accession number GCA\_017355495.1; assembly version: Ogor\_1.0 (odd year); scaffolds: 20,642; contigs: 20,663; N50: 1.8Mb; L50: 386) from NCBI Genome database (<https://www.ncbi.nlm.nih.gov/genome/>) under BioProject PRJNA556728, Fisheries and Oceans Canada (Christensen et al., 2021). Third, we then constructed the FM-index for the reference genome and individually aligned the cleaned, demultiplexed PE reads from each fish to the indexed reference genome using the BWA-MEM algorithm of BWA v.0.7.17 aligner (Li & Durbin, 2009), excluding reads with a minimum quality score of <30. Alignments were sorted, indexed and read pairs were fixed using tools from the SAMtools v.1.9 suite (Li et al., 2009), and we obtained alignment quality control (QC) statistics for the sorted and indexed alignments using BAMQC as implementing in QUALIMAP v.2.2.20 (Okonechnikov et al., 2016). Finally, we used the alignments that passed QC to assemble RAD-loci, call SNPs and construct genotypes for individual fish using the perl script *ref\_map.pl* also implemented in STACKS, which consisted of two components (*gstacks* and *populations*). We ran population analysis in *ref\_map.pl*, with (i) a minimum of 75% of individuals in a population required

to process a RAD-locus for that population (-r 0.75), (ii) a minimum of 70% of populations a RAD-locus must be present in to process a locus (-p 4), (iii) discard unpaired reads (--rm-unpaired-reads), (iv) a minimum mapping quality of 20 to consider a read (--min-mapq 20) and (v) a minimum minor allele count (MAC) of 2 required to process an SNP at a RAD-locus (--min-mac 2).

## 2.2.4 | Data filtering

A recent whole-genome duplication occurred in the common ancestor of salmonids 80million years ago (Mya) following their separation from Esociformes 125 Mya, and a substantial percentage of the salmonid genome remains duplicated (Allendorf et al., 2015; Lien et al., 2016; MacQueen & Johnston, 2014). Therefore, in the genomes of salmonids, there is a prevalence of paralogous loci, which can result in the occurrence of both disomic and tetrasomic inheritance within the same genome (Gidskehaug et al., 2011; McKinney et al., 2017). This implies that the presence of SNPs originating from paralogous regions of the genome, known as multi-site variants (MSVs), can lead to significant distortions in the process of SNP calling and biased population genomics estimates (Nadukkalam Ravindran et al., 2018). Therefore, we base our data filtering on the prior work of McKinney et al. (2017) to filter out multiple copy loci

(duplicated SNPs, MSVs) from the STACKS filtered SNP data set. To achieve this, we used the filtering procedures and custom scripts available in STACKS WORKFLOW v.2 ([https://github.com/enormandeau/stacks\\_workflow](https://github.com/enormandeau/stacks_workflow)). A detailed account of the filtering procedures can be found within the Appendix S1. After filtering out MSVs, we further filtered out SNPs that were located in unplaced scaffolds, that is, contigs that were not part of the 26 chromosomes of the pink salmon reference genome used in the present study. We used PGDSPIDER v.2.1.1.5 (Lischer & Excoffier, 2012) to transform VCF files into input files for the respective downstream population genomics programs, where applicable.

## 2.2.5 | Genome scans

Loci under selection are expected to behave differently from neutral loci in terms of population-related patterns of diversification and are thus inappropriate to infer population demographic history (Holderegger et al., 2006; Luikart et al., 2003). Moreover, neutrality was an assumption underlying the population structure models that we used; thus, we investigated whether SNPs were putatively under selection using a population-level method (BAYESCAN) and an individual-level method (PCADAPT). For BAYESCAN v.2.1 (Foll, 2012; Foll & Gaggiotti, 2008), we set the prior odds ( $pr\_odds$ ) at 10,000 after explorative runs ( $pr\_odds=100$ ; 1000), which is appropriate for the number of markers in our data sets (Foll, 2012; Lotterhos & Whitlock, 2015), ran the model using 10,000 iterations with a thinning interval of 10, a burn-in of 200,000 steps and 20 pilot runs of 5000 iterations. Then, we assessed convergence by visual inspection of trace plots, and additional diagnostic tests were carried out to confirm convergence such as obtaining estimates of autocorrelation and effective sample size, Geweke's convergence diagnostic and Geweke-Brooks plots (Geweke, 1991) with the R-package CODA v.0.19.4 (Plummer et al., 2006). For further post-processing of BAYESCAN output in R, we defined an FDR  $q$ -value threshold of 0.05 and defined the respective locus-specific effects ( $\alpha$ ) and  $q$  combinations to classify SNPs: (i) positive  $\alpha$  and  $q \leq 0.05$  as suggestive of diversifying selection, (ii) negative  $\alpha$  and  $q \geq 0.05$  as suggestive of balancing selection and (iii) positive  $\alpha$  and  $q \geq 0.05$  as suggestive of neutrality. For PCADAPT v.4.3.3 (Luu et al., 2017; Privé et al., 2020), we examined scree plots to determine background population structure as  $K$  principal components and used the first 10 components that captured the majority of population structure in the data. To control the false positive, we set the false discovery rate (FDR) threshold value to 0.05. While BAYESCAN is suitable for our population-level sampling design (Foll, 2012; Foll & Gaggiotti, 2008), PCADAPT is more reliable for species with complex, hierarchical population structure and are less sensitive to admixed individuals and outliers in the data (Luu et al., 2017; Privé et al., 2020). Thus, to derive a putatively neutral SNP data set for estimating population genetics parameters for population structure and gene flow analyses, we removed outliers identified in any of the two methods to be putatively under selection.

## 2.2.6 | Genetic diversity and differentiation

To characterize genetic diversity, we calculated average observed heterozygosity ( $H_o$ ) and within population gene diversity ( $H_s$ ), Wright's inbreeding coefficient ( $F_{IS}$ ) and a population-specific index of differentiation relative to all other sampling populations ( $\beta_{WT}$ , Weir & Goudet, 2017) using the R package HIERFSTAT v.0.5-11 (Goudet, 2005; Goudet & Jombart, 2022). We computed the 95% confidence intervals (CIs) of  $F_{IS}$  and  $\beta_{WT}$  values using the bootstrap method (1000 bootstraps) as implemented in HIERFSTAT. We also calculated the number of private alleles at each locus and sampling population with the function 'privateAlleles' of the R package STRATAG v.2.5.01 (Archer et al., 2017). We calculated genome-wide pairwise  $F_{ST}$  estimates using the Weir and Cockerham (1984) method with the function 'pairwise.WCfst' in HIERFSTAT, and computed the 95% CIs of  $F_{ST}$  values using the bootstrap function 'boot.ppfst' ( $nboot=1000$ ) as implemented in HIERFSTAT. Indices were considered significant if the 95% CI did not include zero.

## 2.2.7 | Genetic clustering analysis

To investigate population clustering, we first computed a matrix of Nei's genetic distance  $D_A$  (Nei, 1972; Nei & Takezaki, 1983) for each pair of populations with the function 'stampNeisD' of the R package STAMPP v.1.6.3 (Pembleton et al., 2013). The distance matrix was then used to perform a principal coordinate analysis (PCoA) using the function 'pcoa' of the R package APE v.5.6.2 (Paradis & Schliep, 2019). Second, we inferred genomic admixture of pink salmon with ADMIXTURE v.1.3 (Shringarpure et al., 2016) implemented within the ADMIXPIPE v.3 pipeline of Musmann et al. (2020). To determine the best-fit number of clusters ( $K$ ) for our data,  $K$  values from 1 to 10 were assessed with 20% cross-validation (CV). Twenty replicates of ADMIXTURE were run at each  $K$ , and the best-fit  $K$  was determined as the value that had the lowest average CV score across replicates. Additionally, given that our sampling was uneven we applied the Puechmaille method (Puechmaille, 2016) to determine the optimum  $K$  as implemented on the STRUCTURESELECTOR web server (<http://lmme.qdio.ac.cn/StructureSelector/>; Li & Liu, 2018). We visualized the ADMIXTURE results using the CLUMPAK server (<http://clumpak.tau.ac.il/>; Kopelman et al., 2015). Last, we assessed the genomic co-ancestry among individuals with FINERADSTRUCTURE (Malinsky et al., 2018). We first inferred a co-ancestry matrix with the script RADpainter. Subsequently, clustering was done with the Markov chain Monte Carlo method of FINERADSTRUCTURE, running for 500,000 generations and sampling every 1000 generations; the first 200,000 generations were discarded as burn-in (non-default parameters: -x 200,000 -y 300,000 -z 1000). We also inferred a tree for visualization with FINERADSTRUCTURE using the tree-building algorithm of Lawson et al. (2012) with 10,000 attempts (non-default parameters: -mT -x 10,000). FINERADSTRUCTURE results were plotted with R scripts included in the FINERADSTRUCTURE package.



## 2.2.8 | Gene flow analysis

To estimate contemporary gene flow among river basins, we employed BAYESASS-SNPs v.1.1 (Mussmann et al., 2019), an expansion of BayesAss v.3.0.4 (Wilson & Rannala, 2003) that facilitates the use of large amounts of SNPs and automates the fine-tuning of model parameters (Mussmann et al., 2019). We performed a first round of analyses with the auto-tuning tool, BA3-SNPs-autotune, developed as part of BA3-SNPs to extensively find optimal mixing parameters for each run with short exploratory runs of BA3-SNPs. We performed a maximum of 10 repetitions, each time discarding 10,000 generations as burn-in and running for 100,000 generations. The number of repetitions required to find optimal mixing parameters was recorded for each, and mixing parameters verified to produce adequate MCMC acceptance rates (i.e.  $0.2 < \text{acceptance rate} < 0.6$ ). Subsequently, we conducted the final analysis by running three replicates for 10,000,000 MCMC generations, sampling every 1000 generations and discarding the first 1000,000 generations as burn-in. Convergence of parameter estimations was checked with CODA.

## 3 | RESULTS

### 3.1 | Species distribution modelling

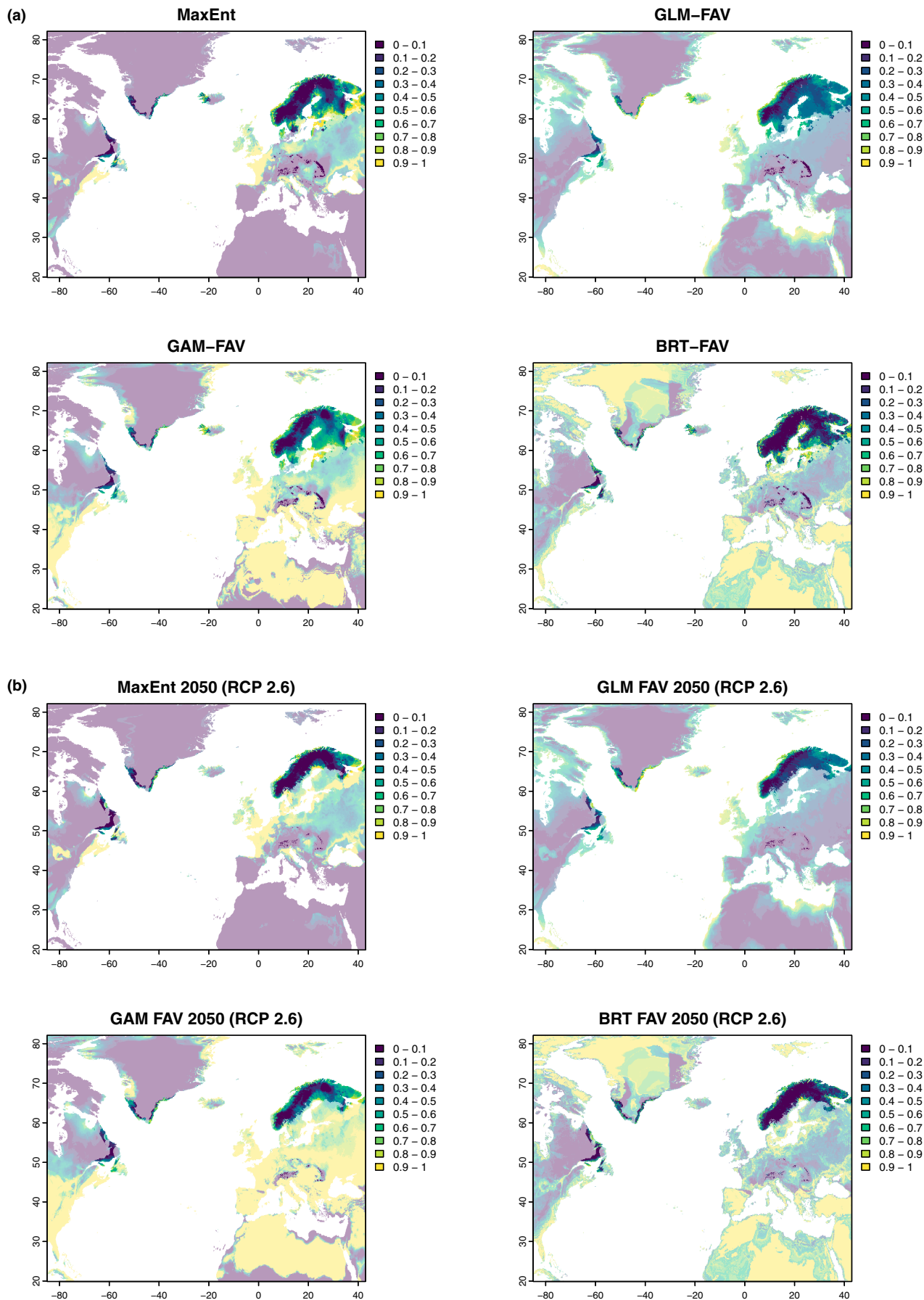
The selected uncorrelated variables (Table 1) included mean diurnal range (BIO2), isothermality (BIO3), maximum temperature of warmest month (BIO5), mean temperature of wettest quarter (BIO8), precipitation seasonality (BIO15), precipitation of warmest quarter (BIO18) and elevation (Elev). The contributions of these variables were largely consistent in the four selected models, and all generated SDMs were accurate and performant, where the cross-validation means of AUC, TSS and Miller's slope across five folds ranged from 0.793 to 0.807, 0.482 to 0.547 and 0.575 to 0.961, respectively (Figure S1). We found that MAXENT suitability predictions provided essentially similar spatial patterns to those of favourability from GLM, GAM and BRT, with which they were highly correlated ( $r^2 > .85$ ; Figure S2). Moreover, quantitative niche comparison with Schoener's  $D$  and Warrant's  $I$  indices revealed high similarity between all SDMs ( $D$ : 0.789–0.912;  $I$ : 0.97–0.964), and between SDMs and the ensemble predictive model ( $D$ : 0.860–0.936;  $I$ : 0.982–0.994), showing that our findings were robust and independent of the chosen modelling algorithm (Table S1). Notably, all SDMs identified the north-eastern most invasion distribution of the pink salmon in Norway as the invasion hotspot, and a relatively high number of suitable habitats for this species in wider Norway (Figure 2, Figure S2). Moreover, favourability remained fairly high across the current invasive distribution range of pink salmon, with model projections across the Arctic and North Atlantic Oceans, including the Baltic and Mediterranean Seas, revealing a high number of suitable habitats (Figure 2), many of which had previously reported pink salmon occurrences, including the Fennoscandian Peninsula,

Greenland, Svalbard, Iceland and the United Kingdom. Furthermore, based on  $D$  and  $I$ , climatically favourable areas inferred from SDM projections remained fairly stable between time periods ( $D$ : 0.789–0.915;  $I$ : 0.949–0.99); however, favourability of several suboptimal areas increased, particularly across the Fennoscandian Peninsula (Figure 2, Table S1).

### 3.2 | Population genomics

We generated a total of 332,338,509 raw paired-end reads from the sequencing of 96 pink salmon samples. Processing of raw Illumina data by the program *process\_radtags* resulted in the removal of 0.06% of the data due to poor sequencing quality, 1.77% of the data due to adapter contamination, 0.82% of the data due to ambiguity in the restriction site and 4.65% of reads due to ambiguous barcodes (inability to attribute a sequencing read to an individual). Overall, we recovered a high percentage of retained reads (92.7%) with a mean of 3.8 million reads per sample. After removing samples that did not sequence well, with <500,000 retained reads, the *gstacks* program generated a catalogue of 638,296 putative RAD loci with a mean insert length of 515.1bp ( $\pm 187.6$ bp) and a mean effective per-sample coverage of 16x (min=5.8x, max=38.0x). After the initial SNP filtering with the *populations* program, we retained 165,231 RAD loci composed of 181,572 SNPs. After secondary filtering, thinning step in STACKS WORKFLOW, purging closely related individuals and removal of MSVs and SNPs within non-chromosomal scaffolds, we retained a final panel of 43,719 polymorphic SNPs (genotyping rate=0.98;  $n=73$ ; Figure S3, Table S2). We detect 9356 putative SNPs under balancing selection and zero SNP putatively under divergent selection with BAYESCAN and 643 putative SNPs under divergent selection with PCADAPT, where 140 outlier SNPs overlapped between the two approaches. We removed all SNPs putatively under selection and retained a final data set of 33,860 putatively neutral SNPs. From this data set, we derived an SNP subset data set of 250 'diagnostic' SNPs with the highest locus-specific  $F_{ST}$  (min=0.097, mean=0.123, max=0.377) with  $MAF > 0.02$  (min=0.021, mean=0.199, max=0.493) to resolve local genetic structure.

The overall patterns of genome-wide diversity based on the 'neutral' full-SNP data set were similar across rivers, with  $H_O$  and  $H_S$  ranging from 0.227 to 0.233 and 0.231 to 0.232, and  $F_{IS}$  ranging from -0.009 to 0.02 (Table 3a). While the  $A_p$  patterns varied, ranging 0–320, where pink salmon in river NDE (Neidenelva) harboured the highest  $A_p$  of 320 alleles and exhibited the highest  $F_{IS}$  of 0.016 (95% CI: 0.014–0.018) (Table 3a). The overall  $\beta_{WT}$  index was shallow (0.001, 95% CI: 0–0.012), where the highest observed  $\beta_{WT}$  within rivers was among pink salmon in the western-most river KKE (Klokkelva; 0.005, 95% CI: 0.001–0.008). With the 250-SNP data set,  $H_O$  and  $H_S$  ranged from 0.23 to 0.272 and 0.233 to 0.268, and  $F_{IS}$  values ranging from -0.058 to -0.065 (Table 3b). The observed  $A_p$  ranged 0–4. The overall  $\beta_{WT}$  index was 0.235 (95% CI: 0.180–0.290), where pink salmon in KKE had a larger  $\beta_{WT}$



**FIGURE 2** Space and time projections of habitat suitability for top performing species distribution models for pink salmon *Oncorhynchus gorbuscha*. Colour scale ranges from darker purple/blue–least suitable–through darker green to brighter yellow/green–most suitable. MAXENT habitat suitability and favourability (FAV) from GLM, GAM and BRT models for space (a) and (b) time projections. Shaded semi-transparent grey overlays depict multivariate environmental similarity surfaces (MESS) maps for environmental dissimilarity of projection layers relative to modelled layers, where predictions in those areas should be treated with strong caution.

**TABLE 3** Genetic variation descriptors of pink salmon within each sampling river basin and fjord, based on the full 3RAD SNP data set (a) and the diagnostic SNP data sets of 250 SNPs based on the highest locus-specific  $F_{ST}$  (b): number of individuals ( $N$ ), number of private alleles ( $A_p$ ), allelic richness ( $A_R$ ), observed heterozygosity ( $H_O$ ), gene diversity ( $H_S$ ), inbreeding coefficient ( $F_{IS}$ ) and sampling population-specific  $F_{ST}$  ( $\beta_{WT}$ ).

River	$N$	$A_p$	$A_R$	$H_O$	$H_S$	$F_{IS}$ (95% CI)	$\beta_{WT}$ (95% CI)
(a) Full SNP data set ( $n=33,860$ )							
KKE	12	0	1.84	0.229	0.231	0.008 (0.004 to 0.012)	0.005 (0.001 to 0.008)
MKE	12	4	1.84	0.227	0.231	0.020 (0.017 to -0.024)	0.003 (-0.001 to 0.006)
NDE	32	320	1.98	0.229	0.232	0.016 (0.014 to 0.018)	-0.003 (-0.005 to -0.000)
JAF	8	4	1.76	0.230	0.231	0.005 (0.000 to 0.010)	0.001 (-0.003 to 0.005)
JFB	9	1	1.78	0.233	0.231	-0.009 (-0.013 to -0.004)	0.000 (-0.003 to 0.004)
Overall	73	65.8	1.84	0.230	0.231	0.008 (0.004 to 0.012)	0.001 (0.000 to 0.003)
(b) Diagnostic SNP data set ( $n=250$ )							
KKE	12	0	1.76	0.245	0.262	0.014 (-0.039 to 0.066)	0.235 (0.180 to 0.290)
MKE	12	0	1.74	0.230	0.233	-0.009 (-0.059 to 0.040)	0.216 (0.158 to 0.275)
NDE	32	0	1.95	0.313	0.296	0.065 (0.035 to 0.096)	0.141 (0.114 to 0.170)
JAF	8	4	1.78	0.241	0.239	-0.058 (-0.105 to -0.003)	0.026 (-0.059 to 0.104)
JFB	9	1	1.75	0.272	0.268	-0.014 (-0.067 to 0.038)	0.120 (0.054 to 0.187)
Overall	73	1	1.79	0.260	0.260	-0.002 (-0.047 to -0.047)	0.148 (0.122 to 0.172)

**TABLE 4** Pairwise  $F_{ST}$  among sampling populations of pink salmon estimated using the full and the diagnostic SNP data sets.

River	KKE	MKE	NDE	JAF	JFB
KKE		0.146 (0.117 to 0.176)	0.054 (0.068 to 0.041)	0.21 (0.189 to 0.231)	0.171 (0.149 to 0.192)
MKE	0.001 (0.001 to 0.002)		0.056 (0.07 to 0.043)	0.195 (0.216 to 0.173)	0.178 (0.153 to 0.199)
NDE	0 (0 to -0.001)	0 (0 to -0.001)		0.117 (0.1 to 0.136)	0.098 (0.081 to 0.115)
JAF	0.003 (0.002 to 0.004)	0.001 (0.002 to 0)	0 (-0.001 to 0)		0.224 (0.199 to 0.247)
JFB	0.001 (0 to 0.002)	0.002 (0.001 to 0.003)	0 (-0.001 to 0)	0.001 (0 to 0.002)	

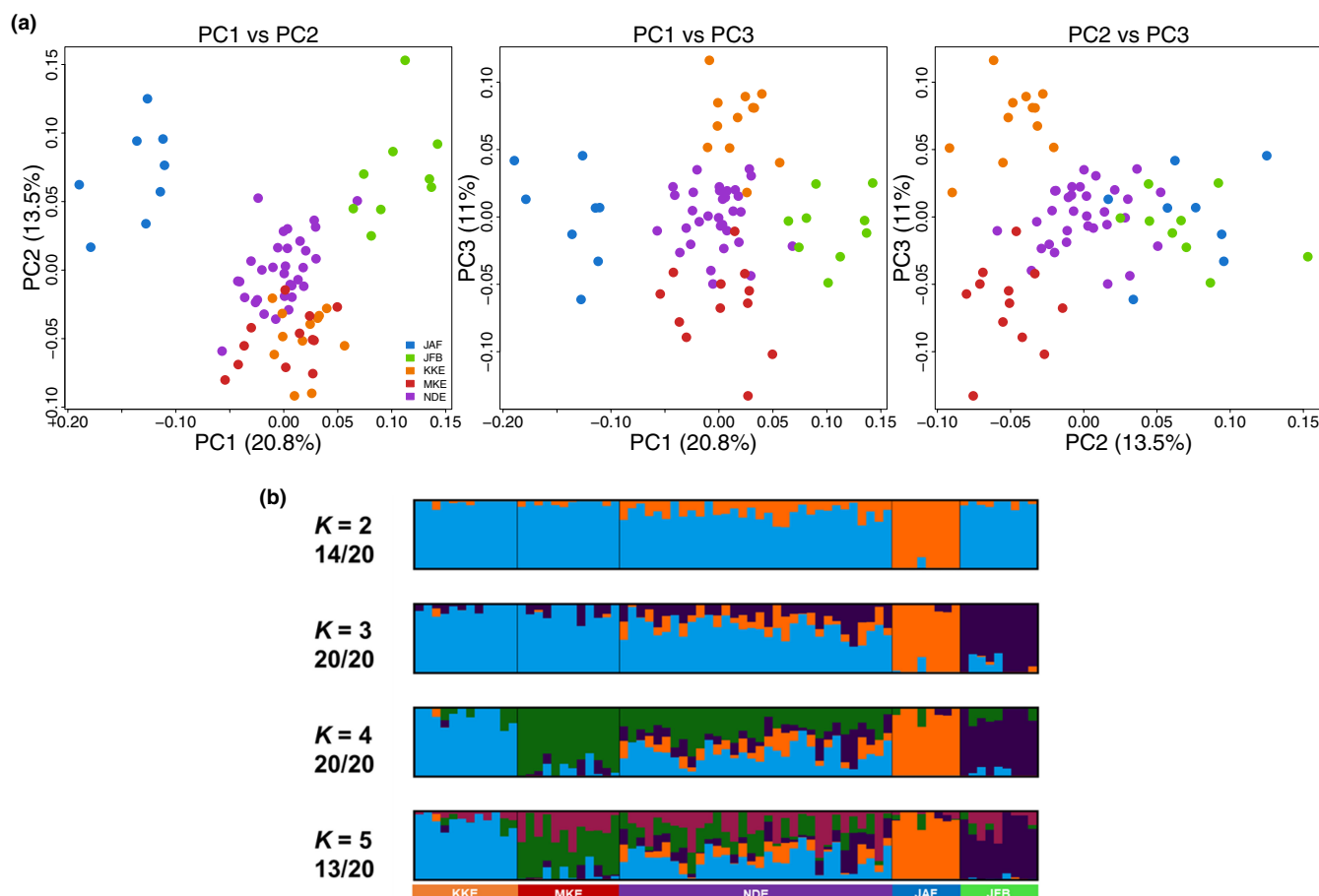
Note:  $F_{ST}$ -values for full data set are below the diagonal and those of the diagnostic data set are above the diagonal.

similar to the full-SNP data sets (Table 3b). Interestingly, there was an increasing trend in  $\beta_{WT}$  from north-east to north-west, where western populations had nearly a twofold higher  $\beta_{WT}$  compared to eastern populations.

The global  $F_{ST}$  for the full-SNP data set was 0.0003, and there was shallow population differentiation with pairwise  $F_{ST}$  ranging from 0.002 to 0.005 and 0 to 0.003 (Table 4). The 250-SNP panel uncovered population differentiation congruent with geographic spatial pattern, with pairwise  $F_{ST}$  ranging from 0.054 to 0.224 (global  $F_{ST}=0.114$ ). We found that pairwise  $F_{ST}$  was lowest among western population comparisons (KKE, MKE and NDE) and highest for the eastern versus western population comparisons and between western populations (JAF and JFB) (Table 4). Interestingly, all pairwise

$F_{ST}$  including NDE was shallow, except for JAF versus NDE pairwise comparison (Table 4).

We explored genetic clustering using several independent methods, and we did not detect genetic structuring with the full SNP data (data not shown) but with the subset data set of 250 diagnostic SNP ('ancestry-informative' SNPs, hereafter). First, the PCoA revealed geographic clustering of pink salmon from different river basins, where samples from respective river basins formed individual clusters except for Neidenelva (NDE). Overall PCoA clustering patterns revealed at least five genetic clusters based on the first three principal components (PCs), which together accounted for 48.7% of the genetic variation (Figure 3a). Second, the post-processing of the ADMIXTURE results based on lowest average CV



**FIGURE 3** Clustering analyses for a global picture of genetic cluster patterns of pink salmon individuals from Varanger Fjord across five river basins based on 250 genome-wide SNPs. (a) Principal coordinate analysis (PCoA) and (b) ADMIXTURE clustering analyses. CLUMPAK-averaged Bayesian clustering (ADMIXTURE) outputs for 20 independent runs of  $K=2-5$  for pink salmon individuals from Finnmark. Numbers below each  $K$ -value show the proportion of runs that converged to solution presented. A single vertical line represents each individual; a black line separates sample sites; the whole sample is divided into  $K$  colours representing the number of clusters inferred. The colours show the estimated individual proportions of cluster membership.

score across replicates identified  $K=3$  as the most likely number of genetic clusters (Figure S4), which was consistent with the number of clusters based on first two principal components and discriminant analysis of the PCoA (Figures 3b). However, the four estimators of the Puechmaillie method (*MedMeaK*, *MaxMeaK*, *MedMedK* and *MaxMedK*) all identified  $K=4$  as the most likely number of genetic clusters (Figure S5). Notably, the genetic composition of NDE was largely composed of admixed pink salmon (Figure 3b). Last, the *FINERADSTRUCTURE* co-ancestry between genotype pairs uncovered clustering of genotypes by river basin, although the NDE genotypes were weakly clustered where the separation into two subgroups was not well supported (Figure 4a). Most individuals from NDE had a genomic admixture profile and corroborated ADMIXTURE analyses.

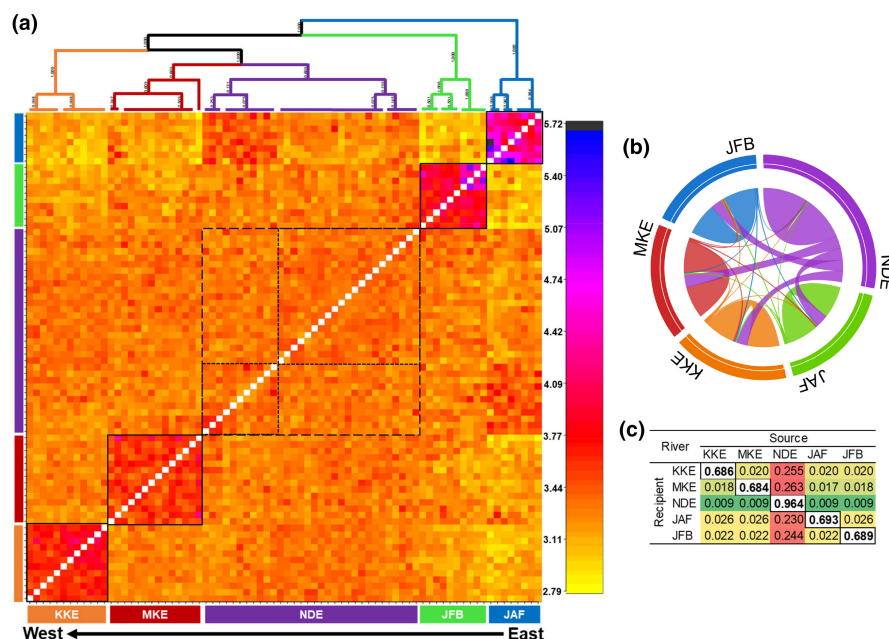
Recent migration rates estimated using the BA3-SNPs program were highly consistent between three independent runs of the program. The average migration rates across the three runs were low and asymmetrical among river basins, except for the high migration rates ( $>0.2$ ) from NDE to all other rivers (Figure 4b,c).

## 4 | DISCUSSION

### 4.1 | Climate favourability and potential invasion hotspots

Based on predictions from our SDMs, we uncovered that Norway harbours a high proportion of climatically favourable areas for the pink salmon, where the northernmost distribution and invasion centre in Finnmark was confirmed as the invasion hotspot. Our models identified that the main environmental drivers for the ecological explosion of the pink salmon were temperature-related (BIO2, BIO3, BIO5 and BIO8) and precipitation-related (BIO15 and BIO18) variables. Additionally, elevation (Elev) was identified as a key driver for habitat suitability, where lower altitudinal areas had the highest favourability. In their non-native range, these main environmental drivers coincide with the seasonal period of arrival and river ascent of adult pink salmon back into their natal rivers for spawning. In Norway, pink salmon migrate up into the rivers primarily from mid-June to late July/early August, where spawning then usually occurs





**FIGURE 4** (a) FINESTRUCTURE pairwise co-ancestry matrix and tree based on 250 genome-wide, diagnostic SNPs. Outer tick marks represent individuals, but labels have been removed and replaced with a coloured bar corresponding to river basin of origin for visualization. The central matrix panel shows co-ancestry between genotypes, with light yellow indicating low co-ancestry, and darker yellows, oranges and reds to black indicating progressively higher co-ancestry. Solid black squares indicate the four major clades corresponding to individual river basins and the direction of invasion-front from east to west. Dashed squares indicate the admixed NDE river basin and the likely two sub-clades. (b) Circular migration plot exhibiting, for each river basin, the fraction of migrant individuals (size of the band) per generation and its origin. (c) The insert table depicts mean migration rate between river basins colour coded from low as green to high as red.

in the latter half of August or the first part of September (Gjelland & Sandlund, 2012; Niemelä et al., 2016; Diaz Pauli et al., 2023; Sandlund et al., 2019). Furthermore, the spawning grounds of the pink salmon are frequently in the lower regions of rivers (~50 km) and within the tidal zone in short coastal streams (Heard, 1991; Sandlund et al., 2019), though this can vary to >100 km upstream (Niemelä et al., 2016; Sandlund et al., 2019 and references therein). Moreover, the projection of our models in space revealed larger favourable areas for pink salmon in the Pan-Atlantic, which corresponded to confirmed occurrences of the species in the Fennoscandian Peninsula (Gordeeva et al., 2015; Niemelä et al., 2016; Staveley & Ahlbeck Bergendahl, 2022; Zubchenko et al., 2004), Svalbard (Hindar et al., 2020; Sandlund et al., 2019), Greenland (Nielsen et al., 2020), Ireland (Millane et al., 2019; Whelan, 2017), Faroe Islands (Eliassen & Johannesen, 2021), Scotland (Armstrong et al., 2018; Skóra, Jones, et al., 2023a) and Iceland (Gudjonsson & Scarnecchia, 2009; Skóra, Guðbergsson, et al., 2023b; Whelan, 2017). This integration of our models with additional information from the primary literature provides high confidence in the robustness of the provided predictions of habitat suitability and climatic favourability for pink salmon in its non-native range. Climate change is expected to accelerate the current spatial invasion-front and expand the species' potential ranges towards southern Europe and the northwest over the next three decades, according to our future projected models, if no effective post-control measures are implemented. Nonetheless, it has not escaped our notice that our models implicitly rely on ecological niche concepts; consequently, we assume that the pink salmon

occurs everywhere where environmental conditions are favourable and that dispersal is not a limiting factor (Jeschke & Strayer, 2008), which does not apply to NIS (Václavík & Meentemeyer, 2009). Indeed, NIS are frequently absent from specific locations not due to poor habitat quality, but because the species has not dispersed to that location due to stochastic events, geographical barriers and dispersal constraints (Araújo et al., 2005; Araújo & Guisan, 2006; Higgins et al., 1999). As a result, we must exercise caution when employing SDM projections in the context of future climate change. Nevertheless, considering the pink salmon's propensity to stray and homing behaviour, along with its widespread distribution beyond its introduced and non-native habitats, we argue that the aforementioned conceptual issue with ecological niche concepts is unlikely to skew our future SDM projections for the species. In fact, all known pink salmon occurrences were strongly supported by our predictions, where the species' introduced range in the White Sea region was well supported in all of our space and time model projections.

## 4.2 | Genetic diversity, genetic structuring and connectivity

Human-mediated introductions, as well as the subsequent establishment and spread of NIS, have the potential to cause a founder or bottleneck effect, resulting in low genetic diversity and an increased likelihood of inbreeding (Colautti et al., 2017; Dlugosch et al., 2015; Estoup et al., 2016; Hänfling, 2007). Strikingly, our results revealed

that genome-wide genetic variation in the 2019 odd-year cohort of non-native populations of the pink salmon were higher than those of native populations in the North Pacific Ocean (Tarpey et al., 2018). With a panel of 16,681 SNPs, Tarpey et al. (2018) found that the mean  $H_O$  and  $H_E$  per population ( $n=7$ ) for the odd-year broodline ranged from 0.154 to 0.172 and 0.156 to 0.172, respectively, with a global  $F_{ST}$  of 0.033 (95% CI: 0.032–0.035). Moreover, we found negligible levels of inbreeding within river basins. Several studies demonstrated a similar trend in that introduced, non-native populations can exhibit a genetic paradox in which they can pass through the population bottleneck with minimal genetic diversity loss, exhibit no negative signs of low genetic diversity or inbreeding, rapidly adapt to their novel environment and successfully establish a self-sustaining population in that novel environment (Bodt et al., 2020; Buchholz et al., 2023; García et al., 2017; Kim et al., 2019; McCann et al., 2014; Wang et al., 2019). Together, these results add to the growing body of knowledge that a founding event is not necessarily linked to a paucity of genetic diversity in invasive species (Gillis et al., 2009; Hänfling, 2007; Roman & Darling, 2007). Furthermore, based on the 'neutral' full-SNP data set, our results likely reflect that pink salmon in Norway may have originated from a single admixed source population. However, the spatial genetic structure inferred with ancestry-informative SNPs uncovered the presence of four ancestral genetic clusters corresponding to river basins, where eastern sampling locations were the most divergent. These genetic clusters are likely the outcome of substantial genetic drift caused by previous successive founder effects during range expansion, combined with the intense culling events in 2017. Furthermore, it has not escaped our notice that the invasive bridgehead effect could be at play here (Bertelsmeier & Keller, 2018; Lombaert et al., 2010). Although we did not have access to contemporary 'native'/introduced populations to explicitly investigate the bridgehead effect in our study system, the contemporary gene flow rates revealed that the unmanaged river network NDE is likely the main source of invaders. Our results did not indicate distance-restricted gene flow based on limited dispersal capability but rather suggested homing behaviour with some degree of straying, where easternmost sampling locations were the most divergent.

### 4.3 | Genome-enabled invasive species management

The management of invasive species requires knowledge of regional-scale dispersal processes and their associated genetic diversity patterns. Results from our population genomic analysis of pink salmon in Varanger Fjord located within the Sør-Varanger Municipality provide actionable information for management officials particularly in terms of major dispersal pathways and the likely source of invaders within the municipality. Genetic clusters aligned fairly well to 'river' basins and to the broad fjord arm divisions of the Varanger Fjord, where these boundaries might provide directly actionable units for municipal management efforts. For instance, river of importance for migratory salmonids within

the Neiden/Munkefjord was found to harbour feral populations of pink salmon with the lowest levels of genetic structure and highest within-population genetic variation. Moreover, our results uncovered that these rivers likely constitute distinct genetic backgrounds and pools of standing genomic variation. Remarkably, a small panel of 11 'ancestry-informative', outlier SNPs revealed pronounced genetic differentiation between early-run and late-run spawners in odd-year lineage of the pink salmon in its native range (Zelenina et al., 2023). These findings supported the hypothesis that the run of pink salmon may be bimodal and driven by temporal separation in the hatching times for offspring of early and late spawners in its native range (Lennox et al., 2023; Taylor, 1980). Interestingly, a similar bimodal run of the odd-year lineage was also reported in several rivers within the invasion centre in Finnmark (Bjerknes, 1977; Lennox et al., 2023). The post-control measures in Finnmark involve the use of river trap systems in the lower parts of rivers and fishing nets (Berntsen et al., 2022), but so far the timing of the establishment and removal of these restricting access measures has often allowed the arrival of both early-run and late-run spawners in the local rivers (personal correspondence with local river guards). Consequently, the current study serves as a benchmark for integrating genomics to answer questions that inform management decisions.

## 5 | CONCLUSIONS

To our knowledge, this study provided the first assessment of ecological niche and species distribution models (SDMs) and population genomics of the pink salmon outside its native distribution. Our integrative framework of SDMs and population genomics provided a general paradigm for understanding climatically favourable areas, population differentiation and dispersal patterns involved in inter-basin straying for the pink salmon in its northernmost invasion and current distribution in Norway. Our study uncovered novel and important results regarding feral population of pink salmon following the 2017 surge and culling events. First, we uncover the invasion hotspots for the pink salmon in Norway and the broader the Pan-Atlantic under current and future climates with SDMs. Second, we found that genome-wide patterns of genetic diversity within populations varied little between river basins but varied with an increasing trend from north-east to north-west when using a subset of 250 ancestry-informative SNPs. Third, our results revealed that the 2019 cohort of the pink salmon constitute of at least five genetic clusters that follow a hierarchical geographic pattern. Last, contemporary migration rates were asymmetrical with Neidenelva (NDE) serving as a source population of invaders in the ongoing secondary spread of the pink salmon in Norway, or at least one of the source areas. These findings derived from our SDMs and population genomic analysis of pink salmon furnish management officials with practical insights, specifically concerning climatically favourable areas, significant dispersal routes and the probable origin of intruders within the

Sør-Varanger Municipality and border Norway. Furthermore, the genome-wide SNP database reported in this study will aid future larger scale studies to shed light on whether the currently implemented post-control measures are inadvertently selecting for invasion-related traits that play a crucial role in successful invasions and distinguish intrinsic genomic features in pink salmon associated with life-history traits that likely evolved within their introduced and native ranges.

## ACKNOWLEDGEMENTS

This work received funding from NIBIO and the Norwegian Ministry of Climate and Environment. We thank all of the local stakeholders that provided genetic samples for our research. The sequencing service was provided by the Norwegian Sequencing Centre ([www.sequencing.uio.no](http://www.sequencing.uio.no)), a national technology platform hosted by the University of Oslo and supported by the 'Functional Genomics' and 'Infrastructure' programmes of the Research Council of Norway and the Southeastern Regional Health Authorities. We also thank the National e-Infrastructure for Research Data (UNINETT Sigma2) for providing high-end computing resources under project number NN9614K.

## CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## PEER REVIEW

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/ddi.13811>.

## DATA AVAILABILITY STATEMENT

Raw, demultiplexed sequencing data are presently accessible at NIBIO-Svanhovd upon request. However, the filtered genome-wide SNP data sets used in the current study are available on Dryad Repository (<https://10.5061/dryad.zw3r228f2>).

## ORCID

Simo Njabulo Maduna  <https://orcid.org/0000-0002-9372-4360>

Paul Eric Aspholm  <https://orcid.org/0000-0002-6352-1781>

Ane-Sofie Bednarczyk Hansen  <https://orcid.org/0009-0005-9616-8303>

Cornelya F. C. Klütsch  <https://orcid.org/0000-0001-8238-2484>

Snorre B. Hagen  <https://orcid.org/0000-0001-8289-7752>

## REFERENCES

- Acevedo, P., & Real, R. (2012). Favourability: Concept, distinctive characteristics and potential usefulness. *Naturwissenschaften*, 99, 515–522.
- Allendorf, F. W., Bassham, S., Cresko, W. A., Limborg, M. T., Seeb, L. W., & Seeb, J. E. (2015). Effects of crossovers between homeologs on inheritance and population genomics in polyploid-derived salmonid fishes. *The Journal of Heredity*, 106, 217–227.
- Allendorf, F. W., & Lundquist, L. L. (2003). Introduction: Population biology, evolution, and control of invasive species. *Conservation Biology*, 17, 24–30.
- Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, 43, 1223–1232.
- Altenritter, M. E., DeBoer, J. A., Maxson, K. A., Casper, A. F., & Lamer, J. T. (2022). Ecosystem responses to aquatic invasive species management: A synthesis of two decades of bigheaded carp suppression in a large river. *Journal of Environmental Management*, 305, 114354.
- Althukov, Y. P., Salmenkova, E. A., & Omelchenko, V. T. (2000). *Salmonid fishes: Population biology, genetics and management*. Blackwell Science.
- Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Araújo, M. B., Anderson, R. P., Barbosa, M. A., Beale, C. M., Dormann, C. F., Early, R., Garcia, R. A., Guisan, A., Maiorano, L., Naimi, B., O'Hara, R. B., Zimmermann, N. E., & Rahbek, C. (2019). Standards for distribution models in biodiversity assessments. *Science Advances*, 5, eaat4858.
- Araújo, M. B., & Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, 33, 1677–1688.
- Araújo, M. B., Pearson, R. G., & Rahbek, C. (2005). Equilibrium of species' distributions with climate. *Ecography*, 28, 693–695.
- Archer, F. I., Adams, P. E., & Schneiders, B. B. (2017). stratag: An R package for manipulating, summarizing and analysing population genetic data. *Molecular Ecology Resources*, 17, 5–11.
- Armstrong, J. D., Bean, C. W., & Wells, A. (2018). The Scottish invasion of pink salmon in 2017. *Journal of Fish Biology*, 93, 8–11.
- Aspinwall, N. (1974). Genetic analysis of North American populations of the pink salmon, *Oncorhynchus gorbuscha*, possible evidence for the neutral mutation-random drift hypothesis. *Evolution; International Journal of Organic Evolution*, 28, 295–305.
- Baquero, R. A., Barbosa, A. M., Ayllón, D., Guerra, C., Sánchez, E., Araújo, M. B., & Nicola, G. G. (2021). Potential distributions of invasive vertebrates in the Iberian Peninsula under projected changes in climate extreme events. *Diversity and Distributions*, 27, 2262–2276.
- Barbet-Massin, M., Rome, Q., Villemant, C., & Courchamp, F. (2018). Can species distribution models really predict the expansion of invasive species? *PLoS One*, 13, e0193085.
- Barbosa, A., Real, R., Muñoz, A. R., & Brown, J. A. (2013). New measures for assessing model equilibrium and prediction mismatch in species distribution models. *Diversity and Distributions*, 19, 1333–1338.
- Barbosa, A. M. (2015). fuzzySim: Applying fuzzy logic to binary similarity indices in ecology. *Methods in Ecology and Evolution*, 6, 853–858.
- Barbosa, A. M., Real, R., & Vargas, J. M. (2009). Transferability of environmental favourability models in geographic space: The case of the Iberian desman (*Galemys pyrenaicus*) in Portugal and Spain. *Ecological Modelling*, 220, 747–754.
- Bax, N., Williamson, A., Aguero, M., Gonzalez, E., & Geeves, W. (2003). Marine invasive alien species: A threat to global biodiversity. *Marine Policy*, 27, 313–323.
- Bayona-Vásquez, N. J., Glenn, T. C., Kieran, T. J., Pierson, T. W., Hoffberg, S. L., Scott, P. A., Bentley, K. E., Finger, J. W., Louha, S., Troendle, N., Diaz-Jaimes, P., Mauricio, R., & Faircloth, B. C. (2019). Adapterama III: Quadruple-indexed, double/triple-enzyme RADseq libraries (2RAD/3RAD). *PeerJ*, 7, e7724.
- Bellard, C., Cassey, P., & Blackburn, T. M. (2016). Alien species as a driver of recent extinctions. *Biology Letters*, 12, 20150623.
- Berntsen, H. H., Sandlund, O. T., Thorstad, E., & Fiske, P. (2020). Pukkellaks i Norge, 2019. *NINA Rapport*, 1821, Norsk institutt for naturforskning. <https://brage.nina.no/nina-xmlui/bitstream/handle/11250/2651741/1821.pdf?sequence=3&isAllowed=y#:~:text=Totalt%20ble%2020025%20pukkellaks%20registrert,tabelle%201%20og%20tabell%202> (In Norwegian).

- Berntsen, H. H., Sandlund, O. T., & Thorstad, E. B. (2022). Pukkellaks i Norge 2021. *NINA Rapport*, 2160, Norsk institutt for naturforskning. <https://brage.nina.no/nina-xmlui/bitstream/handle/11250/3018858/ninarapport2160.pdf?sequence=1&isAllowed=y> (In Norwegian).
- Berntsen, H. H., Sandlund, O. T., Ugedal, O., Thorstad, E. B., Fiske, P., Urdal, K., Skaala, Ø., Fjeldheim, P. T., Skoglund, H., Larsen, B. F., Muladal, R., & Uglem, I. (2018). Pukkellaks i Norge, 2017. *NINA Rapport*, 1571, Norsk institutt for naturforskning. <https://brage.nina.no/nina-xmlui/bitstream/handle/11250/2575646/1571.pdf?sequence=5&isAllowed=y#:~:text=Totalt%20ble%206390%20pukkellaks%20registrert,282%20under%20sportsfisket%20i%20sj%C3%B8en> (In Norwegian).
- Bertelsmeier, C., & Keller, L. (2018). Bridgehead effects and role of adaptive evolution in invasive populations. *Trends in Ecology & Evolution*, 33, 527–534.
- Bjerknes, V. (1977). Evidence of natural production of pink salmon fry *Oncorhynchus gorbuscha* in Finnmark North Norway. *Astarte*, 10, 5–8.
- Bjerknes, V., & Waag, A. B. (1980). Migration and capture of pink salmon, *Oncorhynchus gorbuscha* Walbaum, in Finnmark, North Norway. *Journal of Fish Biology*, 16, 291–297.
- Blackburn, T. M., Pyšek, P., Bacher, S., Carlton, J. T., Duncan, R. P., Jarošík, V., Wilson, J. R. U., & Richardson, D. M. (2011). A proposed unified framework for biological invasions. *Trends in Ecology & Evolution*, 26, 333–339.
- Bodt, L. H., Rollins, L. A., & Zichello, J. M. (2020). Contrasting mitochondrial diversity of European starlings (*Sturnus vulgaris*) across three invasive continental distributions. *Ecology and Evolution*, 10, 10186–10195.
- Bosch, S., & Fernandez, S. (2022). *Sdmpredictors: Species distribution modelling predictor datasets*. R package version 0.2.13. <http://lifewatch.github.io/sdmpredictors/>
- Brenner, R. E., Moffitt, S. D., & Grant, W. S. (2012). Straying of hatchery salmon in Prince William Sound, Alaska. *Environmental Biology of Fishes*, 94, 179–195.
- Buchholz, M. J., Wright, E. A., Grisham, B. A., Bradley, R. D., Arsuffi, T. L., & Conway, W. C. (2023). Low genetic diversity among introduced axis deer: Comments on the genetic paradox and invasive species. *Journal of Mammalogy*, 104, 603–618.
- Burgess, B. T., Irvine, R. L., Howald, G. R., & Russello, M. A. (2021). The promise of genetics and genomics for improving invasive mammal management on islands. *Frontiers in Ecology and Evolution*, 9, 704809.
- Burgess, B. T., Irvine, R. L., & Russello, M. A. (2022). Population genomics of Sitka black-tailed deer supports invasive species management and ecological restoration on islands. *Communications Biology*, 5, 223.
- Bush, A., Mokany, K., Catullo, R., Hoffmann, A., Kellermann, V., Sgrò, C., McEvey, S., & Ferrier, S. (2016). Incorporating evolutionary adaptation in species distribution modelling reduces projected vulnerability to climate change. *Ecology Letters*, 19, 1468–1478.
- Christensen, K. A., Rondeau, E. B., Sakhrani, D., Biagi, C. A., Johnson, H., Joshi, J., Flores, A. M., Leelakumari, S., Moore, R., Pandoh, P. K., Withler, R. E., Beacham, T. D., Leggatt, R. A., Tarpey, C. M., Seeb, L. W., Seeb, J. E., Jones, S. J. M., Devlin, R. H., & Koop, B. F. (2021). The pink salmon genome: Uncovering the genomic consequences of a two-year life cycle. *PLoS One*, 16, e0255752.
- Colautti, R. I., Alexander, J. M., Dlugosch, K. M., Keller, S. R., & Sultan, S. E. (2017). Invasions and extinctions through the looking glass of evolutionary ecology. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 372, 20160031.
- Diaz Pauli, B., Berntsen, H. H., Thorstad, E. B., Homrum, E. I., Lusseau, S. M., Wennevik, V., & Utne, K. R. (2023). Geographic distribution, abundance, diet, and body size of invasive pink salmon (*Oncorhynchus gorbuscha*) in the Norwegian and Barents seas, and in Norwegian rivers. *ICES Journal of Marine Science*, 80, 76–90.
- Dlugosch, K. M., Anderson, S. R., Braasch, J., Cang, F. A., & Gillette, H. D. (2015). The devil is in the details: Genetic variation in introduced populations and its contributions to invasion. *Molecular Ecology*, 24, 2095–2111.
- Dunlop, K. M., Wipfli, M., Muladal, R., & Wierzbinski, G. (2021). Terrestrial and semi-aquatic scavengers on invasive Pacific pink salmon (*Oncorhynchus gorbuscha*) carcasses in a riparian ecosystem in northern Norway. *Biological Invasions*, 23, 973–979.
- Echelle, A. A., & Echelle, A. F. (1997). Genetic introgression of endemic taxa by non-natives: A case study with Leon Springs pupfish and Sheepshead minnow: Introgresión Genética de Taxas Endémicos Debido a Especies No-nativas: Un Caso de Estudio del Pez Cachorro de Leon Springs y el Pez Bolín. *Conservation Biology*, 11, 153–161.
- El-Barougy, R. F., Dakhil, M. A., Halmy, M. W., Gray, S. M., Abdelaal, M., Khedr, A.-H. A., & Bersier, L.-F. (2021). Invasion risk assessment using trait-environment and species distribution modelling techniques in an arid protected area: Towards conservation prioritization. *Ecological Indicators*, 129, 107951.
- Eliassen, K., & Johannesen, U. V. (2021). The increased occurrence of *Oncorhynchus gorbuscha* (Walbaum, 1792) in The Faroe Islands. *BiolInvasions Records*, 10, 390–395.
- Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, 1, 330–342.
- Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77, 802–813.
- Elton, C. S. (1958). *The ecology of invasions by animals and plants*. Methuen.
- Estoup, A., Ravigné, V., Hufbauer, R., Vitalis, R., Gautier, M., & Facon, B. (2016). Is there a genetic paradox of biological invasion? *Annual Review of Ecology, Evolution, and Systematics*, 47, 51–72.
- Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, 32, 3047–3048.
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37, 4302–4315.
- Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24, 38–49.
- Foll, M. (2012). *BayeScan v2.1 user manual*. <http://cmpg.unibe.ch/software/BayeScan/>
- Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180, 977–993.
- Frankham, R. (2005). Resolving the genetic paradox in invasive species. *Heredity*, 94, 385.
- García, K., Melero, Y., Palazón, S., Gosálbez, J., & Castersana, J. (2017). Spatial mixing of mitochondrial lineages and greater genetic diversity in some invasive populations of the American mink (*Neovision vison*) compared to native populations. *Biological Invasions*, 19, 2663–2673.
- Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., Lawrence, D. M., Neale, R. B., Rasch, P. J., Vertenstein, M., Worley, P. H., Yang, Z. L., & Zhang, M. (2011). The community climate system model version 4. *Journal of Climate*, 24, 4973–4991.
- Geweke, J. F. (1991). *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments* (No. 148). Federal Reserve Bank of Minneapolis.
- Giakoumi, S., Katsanevakis, S., Albano, P. G., Azzurro, E., Cardoso, A. C., Cebrian, E., Deidun, A., Edelist, D., Francour, P., Jimenez, C., Mačić, V., Occhipinti-Ambrogi, A., Rilov, G., & Sghaier, Y. R. (2019). Management priorities for marine invasive species. *The Science of the Total Environment*, 688, 976–982.
- Gidskehaug, L., Kent, M., Hayes, B. J., & Lien, S. (2011). Genotype calling and mapping of multisite variants using an Atlantic salmon iSelect SNP array. *Bioinformatics*, 27, 303–310.



- Gillis, N. K., Walters, L. J., Fernandes, F. C., & Hoffman, E. A. (2009). Higher genetic diversity in introduced than in native populations of the mussel *Mytella charruana*: Evidence of population admixture at introduction sites. *Diversity and Distributions*, 15, 784–795.
- Gjelland, K. Ø., & Sandlund, O. T. (2012). Pukkellaks *Oncorhynchus gorbusha*. Norwegian Biodiversity Information Centre, Fact Sheet 283 (In Norwegian), 3 pp.
- Gordeeva, N. V., Salmenkova, E. A., & Prusov, S. V. (2015). Variability of biological and population genetic indices in pink salmon, *Oncorhynchus gorbusha* transplanted into the White Sea basin. *Journal of Ichthyology*, 55, 69–76.
- Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, 5, 184–186.
- Goudet, J., & Jombart, T. (2022). Hierfstat: Estimation and tests of hierarchical F-statistics. <https://www.r-project.org>, <https://github.com/jgx65/hierfstat>
- Graham, C. F., Glenn, T. C., McArthur, A. G., Boreham, D. R., Kieran, T., Lance, S., Manzon, R. G., Martino, J. A., Pierson, T., Rogers, S. M., Wilson, J. Y., & Somers, C. M. (2015). Impacts of degraded DNA on restriction enzyme associated DNA sequencing (RADSeq). *Molecular Ecology Resources*, 15, 1304–1315.
- Greenwell, B., Boehmke, B., Cunningham, J., & GBM Developers. (2022). *gbm: Generalized Boosted Regression Models*. R package version 2.1.8.1. <https://CRAN.R-project.org/package=gbm>
- Gudjonsson, S., & Scarnecchia, D. L. (2009). "Even the evil need a place to live": Wild salmon, salmon farming, and zoning of the Icelandic coastline. *Fisheries*, 34, 477–486.
- Hänfling, B. (2007). Understanding the establishment success of non-indigenous fishes: Lessons from population genetics. *Journal of Fish Biology*, 71, 115–135.
- Hastie, T. J. (2022). *gam: Generalized additive models*. R package version 1.20.2. <https://CRAN.R-project.org/package=gam>
- Hastie, T. J., & Tibshirani, R. J. (1990). *Generalized additive models*. Chapman & Hall.
- Heard, W. R. (1991). Life history of pink salmon (*Oncorhynchus gorbuscha*). In C. Groot & L. Margolis (Eds.), *Pacific salmon life histories* (pp. 121–123). UBC Press.
- Hendry, A. P., Castaic, V., Kinnison, M. T., & Quinn, T. P. (2004). The evolution of philopatry and dispersal: Homing versus straying in salmonids. In A. P. Hendry & S. C. Stearns (Eds.), *Evolution illuminated: Salmon and their relatives* (pp. 52–91). Oxford University Press.
- Higgins, S. I., Richardson, D. M., Cowling, R. M., & Trinder-Smith, T. H. (1999). Predicting the landscape-scale distribution of alien plants and their threat to plant diversity. *Conservation Biology*, 13, 303–313.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25, 1965–1978.
- Hijmans, R. J., Phillips, S., Leathwick, J., & Elith, J. (2022). *Dismo: Species distribution modeling*. R package version 1.3-9. <https://CRAN.R-project.org/package=dismo>
- Hindar, K., Hole, L. R., Kausrud, K. L., Malmstrøm, M., Rimstad, E., Robertson, L., Sandlund, O. T., Thorstad, E. B., Vollset, K., de Boer, H., & Eldegard, K. (2020). Assessment of the risk to Norwegian biodiversity and aquaculture from pink salmon (*Oncorhynchus gorbuscha*). VKM Report, 01, 2020 (In Norwegian).
- Hoffberg, S. L., Kieran, T. J., Catchen, J. M., Devault, A., Faircloth, B. C., Mauricio, R., & Glenn, T. C. (2016). RADcap: Sequence capture of dual-digest RADseq libraries with identifiable duplicates and reduced missing data. *Molecular Ecology Resources*, 16, 1264–1278.
- Hogg, C. J., Ottewell, K., Latch, P., Rossetto, M., Biggs, J., Gilbert, A., Richmond, S., & Belov, K. (2022). Threatened species initiative: Empowering conservation action using genomic resources. *Proceedings of the National Academy of Sciences of the United States of America*, 119, e2115643118.
- Holderegger, R., Kamm, U., & Gugerli, F. (2006). Adaptive vs. neutral genetic diversity: Implications for landscape genetics. *Landscape Ecology*, 21, 797–807.
- Husa, V., Agnalt, A.-L., Berntsen, H., Falkenhaus, T., Fossøy, F., Forsgren, E., Grefsrud, E. S., Hjelset, A. M., Hanssen, F., Husby, E., Jelmert, A., Mortensen, S., Olsen, S. A., & Sandvik, H. (2022). *Alien marine species in Norway— Mapping, monitoring and assessment of vectors for introductions*. Rapport Fra Havforskningen 2022-8. <https://www.hi.no/hi/nettrapporter/rapport-fra-havforskningen-en-2022-8>
- Jeschke, J. M., & Strayer, D. L. (2008). Usefulness of bioclimatic models for studying climate change and invasive species. *Annals of the New York Academy of Sciences*, 1134, 1–24.
- Jonsson, N., & Jonsson, B. (2018). Comment on "The pink salmon invasion: A Norwegian perspective" (Mo et al., 2018). *Journal of Fish Biology*, 93, 584–585.
- Kim, I. R., Choi, W., Kim, A., Lim, J., Lee, D. H., & Lee, J. R. (2019). Genetic diversity and population structure of nutria (*Myocastor coypus*) in South Korea. *Animals*, 9, 1164.
- Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, 15, 1179–1191.
- Lake, T. A., Briscoe Runquist, R. D., & Moeller, D. A. (2020). Predicting range expansion of invasive species: Pitfalls and best practices for obtaining biologically realistic projections. *Diversity and Distributions*, 26, 1767–1779.
- Lawson, D. J., Hellenthal, G., Myers, S., & Falush, D. (2012). Inference of population structure using dense haplotype data. *PLoS Genetics*, 8, e1002453.
- Legendre, P., & Legendre, L. (2012). *Numerical ecology* (3rd ed.). Elsevier.
- Lennox, R. J., Berntsen, H. H., Garseth, Å. H., Hinch, S. G., Hindar, K., Ugedal, O., Utne, K. R., Vollset, K. W., Whoriskey, F. G., & Thorstad, E. B. (2023). Prospects for the future of pink salmon in three oceans: From the native Pacific to the novel Arctic and Atlantic. *Fish and Fisheries*, 24, 759–776.
- Letnic, M., Webb, J. K., Jessop, T. S., & Dempster, T. (2015). Restricting access to invasion hubs enables sustained control of an invasive vertebrate. *Journal of Applied Ecology*, 52, 341–347.
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 25, 1754–1760.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25, 2078–2079.
- Li, Y. L., & Liu, J. X. (2018). StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Molecular Ecology Resources*, 18, 176–177.
- Lien, S., Koop, B. F., Sandve, S. R., Miller, J. R., Kent, M. P., Nome, T., Hvidsten, T. R., Leong, J. S., Minkley, D. R., Zimin, A., Grammes, F., Grove, H., Gjuvsland, A., Walenz, B., Hermansen, R. A., von Schalburg, K., Rondeau, E. B., di Genova, A., Samy, J. K. A., ... Davidson, W. S. (2016). The Atlantic salmon genome provides insights into rediploidization. *Nature*, 533, 200–205.
- Limborg, M. T., Waples, R. K., Seeb, J. E., & Seeb, L. W. (2014). Temporally isolated lineages of pink salmon reveal unique signatures of selection on distinct pools of standing genetic variation. *The Journal of Heredity*, 105, 741–751.
- Lischer, H. E., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics (Oxford, England)*, 28(2), 298–299.
- Lombaert, E., Guillemaud, T., Cornuet, J. M., Malausa, T., Facon, B., & Estoup, A. (2010). Bridgehead effect in the worldwide invasion of the biocontrol harlequin ladybird. *PLoS One*, 5, e9743.

- Lotterhos, K. E., & Whitlock, M. C. (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, 24, 1031–1046.
- Luikart, G., England, P. R., Tallmon, D., Jordan, S., & Taberlet, P. (2003). The power and promise of population genomics: From genotyping to genome typing. *Nature Reviews. Genetics*, 4, 981–994.
- Lurgi, M., Wells, K., Kennedy, M., Campbell, S., & Fordham, D. A. (2016). A landscape approach to invasive species management. *PLoS One*, 11, e0160417.
- Luu, K., Bazin, E., & Blum, M. G. (2017). Pcadapt: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, 17, 67–77.
- Macqueen, D. J., & Johnston, I. A. (2014). A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. *Proceedings of the Royal Society B: Biological Sciences*, 281, 20132881.
- Malinsky, M., Trucchi, E., Lawson, D. J., & Falush, D. (2018). RADpainter and fineRADstructure: Population inference from RADseq data. *Molecular Biology and Evolution*, 35, 1284–1290.
- Mansfield, E. R., & Helms, B. P. (1982). Detecting multicollinearity. *The American Statistician*, 36, 158–160.
- Marquardt, D. W. (1970). Generalized inverses, ridge regression, biased linear estimation and nonlinear estimation. *Technometrics*, 12, 591–612.
- Matheson, P., & McGaughan, A. (2022). Genomic data is missing for many highly invasive species, restricting our preparedness for escalating incursion rates. *Scientific Reports*, 12, 13987.
- Mathieu-Bégné, E., Loot, G., Mazé-Guilmo, E., Mullet, V., Genthon, C., & Blanchet, S. (2021). Combining species distribution models and population genomics underlines the determinants of range limitation in an emerging parasite. *Ecography*, 44, 307–319.
- McCann, B. E., Malek, M. J., Newman, R. A., Schmit, B. S., Swafford, S. R., Sweitzer, R. A., & Simmons, R. B. (2014). Mitochondrial diversity supports multiple origins for invasive pigs. *The Journal of Wildlife Management*, 78, 202–213.
- McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (2nd ed.). Chapman & Hall.
- McKinney, G. J., Waples, R. K., Seeb, L. W., & Seeb, J. E. (2017). Paralogs are revealed by proportion of heterozygotes and deviations in read ratios in genotyping-by-sequencing data from natural populations. *Molecular Ecology Resources*, 17, 656–669.
- Mellin, C., Lurgi, M., Matthews, S., MacNeil, M. A., Caley, M. J., Bax, N., Przeslawski, R., & Fordham, D. A. (2016). Forecasting marine invasions under climate change: Biotic interactions and demographic processes matter. *Biological Conservation*, 204, 459–467.
- Melo-Merino, S. M., Reyes-Bonilla, H., & Lira-Noriega, A. (2020). Ecological niche models and species distribution models in marine environments: A literature review and spatial analysis of evidence. *Ecological Modelling*, 415, 108837.
- Millane, M., Walsh, L., Roche, W. K., & Gargan, P. G. (2019). Unprecedented widespread occurrence of pink salmon *Oncorhynchus gorbuscha* in Ireland in 2017. *Journal of Fish Biology*, 95, 651–654.
- Miller, M. E., Hui, S. L., & Tierney, W. M. (1991). Validation techniques for logistic regression models. *Statistics in Medicine*, 10, 1213–1226.
- Mo, T. A., Thorstad, E. B., Sandlund, O. T., Berntsen, H. H., Fiske, P., & Uglem, I. (2018). The pink salmon invasion: A Norwegian perspective. *Journal of Fish Biology*, 93, 5–7.
- Mrnak, J. T., Sikora, L. W., Zanden, M. J. V., & Sass, G. G. (2023). Applying panarchy theory to aquatic invasive species management: A case study on invasive rainbow smelt *Osmerus mordax*. *Reviews in Fisheries Science & Aquaculture*, 31, 66–85.
- Muñoz, A.-R., & Real, R. (2006). Assessing the potential range expansion of the exotic monk parakeet in Spain. *Diversity and Distributions*, 12, 656–665.
- Mussmann, S. M., Douglas, M. R., Chafin, T. K., & Douglas, M. E. (2019). BA3-SNPs: Contemporary migration reconfigured in BayesAss for next-generation sequence data. *Methods in Ecology and Evolution*, 10, 1808–1813.
- Mussmann, S. M., Douglas, M. R., Chafin, T. K., & Douglas, M. E. (2020). ADMIXPIPE: Population analyses in ADMIXTURE for non-model organisms. *BMC Bioinformatics*, 21, 337.
- Nadukkalam Ravindran, P., Bentzen, P., Bradbury, I. R., & Beiko, R. G. (2018). PMERGE: Computational filtering of paralogous sequences from RAD-seq data. *Ecology and Evolution*, 8, 7002–7013.
- Nei, M. (1972). Genetic distance between populations. *American Naturalist*, 106, 283–292.
- Nei, M., & Takezaki, N. (1983). Estimation of genetic distances and phylogenetic trees from DNA analysis. *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production*, 21, 405–412.
- Nielsen, J., Rosing-Asvid, A., Meire, L., & Nygaard, R. (2020). Widespread occurrence of pink salmon (*Oncorhynchus gorbuscha*) throughout Greenland coastal waters. *Journal of Fish Biology*, 96, 1505–1507.
- Niemelä, E., Johansen, N., Zubchenko, A. V., Dempson, J. B., Veselov, A., Ieshko, E. P., Barskaya, Y., Novokhatskaya, O. V., Shulman, B. S., Lämsman, M., Hassinen, E., Kuusela, J., Haantie, J., Kylmäaho, M., Kivilahti, E., Arvola, K.-M., & Kalske, T. H. (2016). Pink salmon in the Barents region. Office of the Finnmark County Governor Department of Environmental Affairs. Report 3. [https://www.suomenkalakirjasto.fi/wp-content/uploads/2016/04/Pink-salmon\\_19.08.2016.compressed.pdf](https://www.suomenkalakirjasto.fi/wp-content/uploads/2016/04/Pink-salmon_19.08.2016.compressed.pdf)
- Okonechnikov, K., Conesa, A., & García-Alcalde, F. (2016). Qualimap 2: Advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics*, 32, 292–294.
- Paradis, E., & Schliep, K. (2019). Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35, 526–528.
- Pearce, J., & Ferrier, S. (2000). Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, 133, 225–245.
- Pembleton, L. W., Cogan, N. O., & Forster, J. W. (2013). StAMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources*, 13, 946–952.
- Pereira, P. F., Barbosa, A. M., Godinho, C., Salgueiro, P. A., Silva, R. R., & Lourenço, R. (2020). The spread of the red-billed leiothrix (*Leiothrix lutea*) in Europe: The conquest by an overlooked invader? *Biological Invasions*, 22, 709–722.
- Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*, 7, e37135.
- Phillips, S. (2021). *Maxnet: Fitting 'Maxent' species distribution models with 'glmnet'*. R package version 0.1.4. <https://CRAN.R-project.org/package=maxnet>
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190, 231–259.
- Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News*, 6, 7–11.
- Privé, F., Luu, K., Vilhjálmsson, B. J., & Blum, M. G. B. (2020). Performing highly efficient genome scans for local adaptation with R package pcadapt version 4. *Molecular Biology and Evolution*, 37, 2153–2154.
- Puechmaille, S. J. (2016). The program structure does not reliably recover the correct population structure when sampling is uneven: Subsampling and new estimators alleviate the problem. *Molecular Ecology Resources*, 16, 608–627.

- Quinn, T. P. (2005). *The behavior and ecology of pacific salmon and trout*. American Fisheries Society/University of Washington Press.
- R Core Team. (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Radchenko, V. I., Beamish, R. J., Heard, W. R., & Temnykh, O. S. (2018). Ocean ecology of pink salmon. In R. J. Beamish (Ed.), *The ocean ecology of Pacific salmon and trout* (pp. 15–160). American Fisheries Society.
- Rapacciolo, G., Roy, D. B., Gillings, S., Fox, R., Walker, K., & Purvis, A. (2012). Climatic associations of British species distributions show good transferability in time but low predictive accuracy for range change. *PLoS One*, 7, e40212.
- Real, R., Barbosa, A. M., & Bull, J. W. (2017). Species distributions, quantum theory, and the enhancement of biodiversity measures. *Systematic Biology*, 66, 453–462.
- Real, R., Barbosa, A. M., & Vargas, J. M. (2006). Obtaining environmental favourability functions from logistic regression. *Environmental and Ecological Statistics*, 13, 237–245.
- Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, 28, 4737–4754.
- Rödger, D., Schmidtlein, S., Veith, M., & Lötters, S. (2009). Alien invasive slider turtle in unpredicted habitat: A matter of niche shift or of predictors studied? *PLoS One*, 4, e7843.
- Roman, J., & Darling, J. A. (2007). Paradox lost: Genetic diversity and the success of aquatic invasions. *Trends in Ecology & Evolution*, 22, 454–464.
- Sakai, A. K., Allendorf, F. W., Holt, J. S., Lodge, D. M., Molofsky, J., With, K. A., Baughman, S., Cabin, R. J., Cohen, J. E., Ellstrand, N. C., McCauley, D. E., O'Neil, P., Parker, I. M., Thompson, J. N., & Weller, S. G. (2001). The population biology of invasive species. *Annual Review of Ecology and Systematics*, 32, 305–332.
- Sandlund, O. T., Berntsen, H. H., Fiske, P., Kuusela, J., Muladal, R., Niemelä, E., Uglem, I., Forseth, T., Mo, T. A., Thorstad, E. B., Veselov, A. E., Vollset, K. W., & Zubchenko, A. V. (2019). Pink salmon in Norway: The reluctant invader. *Biological Invasions*, 21, 1033–1054.
- Schoener, T. W. (1968). Anolis lizards of Bimini: Resource partitioning in a complex fauna. *Ecology*, 49, 704–726.
- Scoble, J., & Lowe, A. J. (2010). A case for incorporating phylogeography and landscape genetics into species distribution modelling approaches to improve climate adaptation and conservation planning. *Diversity and Distributions*, 16, 343–353.
- Scott, W. B., & Crossman, E. J. (1973). Freshwater fishes of Canada. *Fisheries Research Board of Canada Bulletin*, 184, 1–966.
- Shackleton, R. T., Richardson, D. M., Shackleton, C. M., Bennett, B., Crowley, S. L., Dehnen-Schmutz, K., Estévez, R. A., Fischer, A., Kueffer, C., Kull, C. A., Marchante, E., Novoa, A., Potgieter, L. J., Vaas, J., Vaz, A. S., & Larson, B. M. H. (2019). Explaining people's perceptions of invasive alien species: A conceptual framework. *Journal of Environmental Management*, 229, 10–26.
- Shringarpure, S. S., Bustamante, C. D., Lange, K., & Alexander, D. H. (2016). Efficient analysis of large datasets and sex bias with ADMIXTURE. *BMC Bioinformatics*, 17, 218.
- Sillero, N., & Barbosa, A. M. (2021). Common mistakes in ecological niche models. *International Journal of Geographical Information Science*, 35, 213–226.
- Simberloff, D., Martin, J. L., Genovesi, P., Maris, V., Wardle, D. A., Aronson, J., Courchamp, F., Galil, B., García-Berthou, E., Pascal, M., Pyšek, P., Sousa, R., Tabacchi, E., & Vilà, M. (2013). Impacts of biological invasions: What's what and the way forward. *Trends in Ecology & Evolution*, 28, 58–66.
- Simberloff, D., Parker, I. M., & Windle, P. N. (2005). Introduced species policy, management, and future research needs. *Frontiers in Ecology and the Environment*, 3, 12–20.
- Skóra, M. E., Guðbergsson, G., Copp, G. H., & Jones, J. I. (2023b). Evidence of successful recruitment of non-native pink salmon *Oncorhynchus gorbuscha* in Iceland. *Journal of Fish Biology*, 102, 721–726. <https://doi.org/10.1111/jfb.15556>
- Skóra, M. E., Jones, J. I., Youngson, A. F., Robertson, S., Wells, A., Lauridsen, R. B., & Copp, G. H. (2023a). Evidence of potential establishment of pink salmon *Oncorhynchus gorbuscha* in Scotland. *Journal of Fish Biology*, 102, 721–726.
- Staveley, T. A. B., & Ahlbeck Bergendahl, I. (2022). Pink salmon distribution in Sweden: The calm before the storm? *Ecology and Evolution*, 12, e9194.
- Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, 240, 1285–1293.
- Tarpey, C. M., Seeb, J. E., McKinney, G. J., Templin, W. D., Bugaev, A., Sato, S., & Seeb, L. W. (2018). Single-nucleotide polymorphism data describe contemporary population structure and diversity in allochronic lineages of pink salmon (*Oncorhynchus gorbuscha*). *Canadian Journal of Fisheries and Aquatic Sciences*, 75(6), 987–997.
- Taylor, S. G. (1980). Marine survival of pink salmon fry from early and late spawners. *Transactions of the American Fisheries Society*, 109, 79–82.
- Thirstrup, J. P., Ruiz-Gonzalez, A., Pujolar, J. M., Larsen, P. F., Jensen, J., Randi, E., Zalewski, A., & Pertoldi, C. (2015). Population genetic structure in farm and feral American mink (*Neovison vison*) inferred from RAD sequencing-generated single nucleotide polymorphisms. *Journal of Animal Science*, 93, 3773–3782.
- Ueda, H. (2012). Physiological mechanisms of imprinting and homing migration in Pacific salmon *Oncorhynchus* spp. *Journal of Fish Biology*, 81, 543–558.
- Václavík, T., & Meentemeyer, R. K. (2009). Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? *Ecological Modelling*, 220, 3248–3258.
- Valavi, R., Elith, J., Lahoz-Monfort, J. J., & Guillera-Aroita, G. (2018). blockCV: An R package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. *Biorxiv*, 357798.
- Wang, J., Gaughan, S., Lamer, J. T., Deng, C., Hu, W., Wachholtz, M., Qin, S., Nie, H., Liao, X., Ling, Q., Li, W., Zhu, L., Bernatchez, L., Wang, C., & Lu, G. (2019). Resolving the genetic paradox of invasions: Preadapted genomes and postintroduction hybridization of bigheaded carps in the Mississippi River basin. *Evolutionary Applications*, 13, 263–277.
- Warren, D. L., Glor, R. E., & Turelli, M. (2008). Environmental niche equivalency versus conservatism: Quantitative approaches to niche evolution. *Evolution; International Journal of Organic Evolution*, 62, 2868–2883.
- Warren, D. L., Glor, R. E., & Turelli, M. (2010). ENMTools: A toolbox for comparative studies of environmental niche models. *Ecography*, 33, 607–611.
- Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 38, 1358–1370.
- Weir, B. S., & Goudet, J. (2017). A unified characterization of population structure and relatedness. *Genetics*, 206, 2085–2103.
- Werkowska, W., Márquez, A. L., Real, R., & Acevedo, P. (2017). A practical overview of transferability in species distribution modeling. *Environmental Reviews*, 133, 127–133.
- Whelan, K. (2017). Pink invaders. *Off the scale*, issue 18, Sept–Oct 2017, pp. 14–21. [http://www.offthescaleangling.ie/issue18/offthescale\\_issue18.pdf](http://www.offthescaleangling.ie/issue18/offthescale_issue18.pdf)
- Wilson, G. A., & Rannala, B. (2003). Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*, 163, 1177–1191.
- Yates, K. L., Bouchet, P. J., Caley, M. J., Mengersen, K., Randin, C. F., Parnell, S., Fielding, A. H., Bamford, A. J., Ban, S., Barbosa, A. M., Dormann, C. F., Elith, J., Embling, C. B., Ervin, G. N., Fisher, R., Gould, S., Graf, R. F., Gregr, E. J., Halpin, P. N., ... Sequeira, A. M. M.

- (2018). Outstanding challenges in the transferability of ecological models. *Trends in Ecology & Evolution*, 33, 790–802.
- Zelenina, D. A., Sergeev, A. A., Kositsina, A. I., Soshnina, V. A., & Shpigalskaya, N. Y. (2023). SNP-based discrimination of pink salmon stocks of the Sea of Okhotsk basin: Resolution of the approach and possible ways to increase it. *Frontiers in Marine Science*, 10, 1140538.
- Zhang, Z., Mammola, S., McLay, C. L., Capinha, C., & Yokota, M. (2020). To invade or not to invade? Exploring the niche-based processes underlying the failure of a biological invasion using the invasive Chinese mitten crab. *The Science of the Total Environment*, 728, 138815.
- Zubchenko, A. V., Veselov, A. E., & Kalyuzhin, S. (2004). *Pink Salmon (Oncorhynchus gorbuscha): Problems in acclimatization in Europe, North Russia*. Polar research Institute of Marine Fisheries and Oceanography. Murmansk: PINRO (In Russian).
- Zurell, D., Franklin, J., König, C., Bouchet, P. J., Dormann, C. F., Elith, J., Fandos, G., Feng, X., Guiller-Aroita, G., Guisan, A., Lahoz-Monfort, J. J., Leitão, P. J., Park, D. S., Peterson, A. T., Rapacciuolo, G., Schmatz, D. R., Schröder, B., Serra-Diaz, J. M., Thuiller, W., ... Merow, C. (2020). A standard protocol for reporting species distribution models. *Ecography*, 43, 1261–1277.

### BIOSKETCH

The participant team comprises a heterogeneous group of experts in conservation genetics, molecular ecology and ecology from the Department of Ecosystems in the Barents Region at the NIBIO Svanhovd Research Station in Norway. The department engages in research and methodological advancements in the domains of ecology and ecological genetics, with a specific emphasis on species, populations, ecosystems and ecosystem services in the Barents region. The department actively explores

a range of inquiries pertaining to sustainability and innovation in research areas including biodiversity, climate change, conservation biology, pollution, protected areas, large carnivores and strategies for sustainable living and resource utilization.

**Author contributions:** The study was conceived and designed by SNM and SBH, who also wrote the original draft of the manuscript. PEA and SBH assisted in the gathering of genetic samples in the field. SNM designed and performed the experiments with assistance from ABH. SNM performed the ecoinformatics and bioinformatics analyses. SNM spearheaded the writing process, with editorial contributions from PEA, CFCK and SBH. The submission has been approved by all relevant authors.

### SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Maduna, S. N., Aspholm, P. E., Hansen, A.-S., Klütsch, C. F. C., & Hagen, S. B. (2024). Ecological niche modelling and population genomics provide insights into the geographic and demographic 'explosion' of a non-indigenous salmonid. *Diversity and Distributions*, 30, e13811. <https://doi.org/10.1111/ddi.13811>