

Kolarctic CBC – Project KO4178; Conserving our Atlantic salmon as a sustainable resource for people in the North; fisheries and conservation in the context of growing threats and a changing environment.

Report XX. Application of the updated genetic baseline for genetic stock identification of Atlantic salmon in commercial fisheries in northern Norway

Ozerov M.Yu.^{1,2}, Wennevik V.³, Niemelä E.^{1,4,5}, Vasemägi A.^{6,7}, Prusov S.V.⁸, Vorontsova T.¹, Kalske T.⁴, Olsen A.A.⁴, Høstmark M.S.⁴, Smeland A.F.⁴

¹*Biodiversity Unit, University of Turku, 20014 Turku, Finland*

²*Department of Biology, University of Turku, 20014 Turku, Finland*

³*Department of Aquaculture, Institute of Marine Research, Nordnes, PO Box 1870, 5817 Bergen, Norway*

⁴*County Governor of Troms and Finnmark, 9800 Vadsø, Norway*

⁵*Tmi Olli van der Meer, Hiimonkatu 14, 90850 Haukipudas, Finland*

⁶*Department of Aquatic Resources, Institute of Freshwater Research, Swedish University of Agricultural Sciences, 17893 Drottningholm, Sweden*

⁷*Chair of Aquaculture, Institute of Veterinary Medicine and Animal Sciences, Estonian University of Life Sciences, 46A Kreutzwaldi St., 51006 Tartu, Estonia*

⁸*The Polar branch of Russian Federal Research Institute of Fisheries and Oceanography “VNIRO” (“PINRO” named after N.M.Knipovich, Murmansk, Russia)*



Co-funded by the
European Union

Corresponding author: Mikhail Ozerov, Biodiversity Unit, 20014 University of Turku, Finland, tel. +358 29 450 4241, e-mail: mikoze@utu.fi; mikhail.ozerov@gmail.com.

Statsforvalteren i Troms og Finnmark PB 700, 9815 VADSØ www.statsforvalteren/troms-finnmark/	ISBN 978-82-94021-20-8
CoASal – KO4178 project	Date: 31.1.2023

Report XX. Application of the updated genetic baseline for genetic stock identification of Atlantic salmon in commercial fisheries in northern Norway

Ozerov M.Yu.^{1,2}, Wennevik V.³, Niemelä E.^{1,4,5}, Vasemägi A.^{6,7}, Prusov S.V.⁸, Vorontsova T.¹, Kalske T.⁴, Olsen A.A.⁴, Høstmark M.S.⁴, Smeland A.F.⁴

¹*Biodiversity Unit, University of Turku, 20014 Turku, Finland*

²*Department of Biology, University of Turku, 20014 Turku, Finland*

³*Department of Aquaculture, Institute of Marine Research, Nordnes, PO Box 1870, 5817 Bergen, Norway*

⁴*County Governor of Troms and Finnmark, 9800 Vadsø, Norway*

⁵*Tmi Olli van der Meer, Hiomonkatu 14, 90850 Haukipudas, Finland*

⁶*Department of Aquatic Resources, Institute of Freshwater Research, Swedish University of Agricultural Sciences, 17893 Drottningholm, Sweden*

⁷*Chair of Aquaculture, Institute of Veterinary Medicine and Animal Sciences, Estonian University of Life Sciences, 46A Kreutzwaldi St., 51006 Tartu, Estonia*

⁸*The Polar branch of Russian Federal Research Institute of Fisheries and Oceanography “VNIRO” (“PINRO” named after N.M.Knipovich, Murmansk, Russia)*

Abstract

Atlantic salmon is an important exploitable resource for people living both along the rivers and the coast of northern Europe. Sustainable management of Atlantic salmon presumes preservation of many populations while at the same time ensuring that salmon remains available both for commercial and recreational fisheries to the benefit of both local people and visitors. However, fishery is one of the major factors affecting the production and survival of salmon populations. Moreover, coastal and riverine fisheries exploit salmon along its migration routes and the fisheries are considered to be mixed-stock fisheries, as salmon from different populations may contribute to the catches, challenging management. Application of genetic stock identification (GSI) methods allows for obtaining the knowledge of the level of contribution of different salmon populations to the catches, which in turn enables to infer exploitation rates of various stocks by the fishery.

In this work, by applying GSI, we aimed to investigate whether exploitation patterns of salmon populations in northern Norway have changed since the previous years. We focused on Varanger fjord area (Sør-Varanger, Nesseby and Vadsø municipalities), which can be considered as a reference area, given that interceptory fishery here has been the most intensive both for salmon originating from the rivers of Finnmark in Norway and from the Kola Peninsula in Russia. In addition, we compared recent and previous catch compositions in other Finnmark municipalities (spanning from Vardø to Alta). Further, as the management regime for coastal fisheries was adjusted in the region based on the findings of Kolarctic Salmon project, we aimed to explore if this have had an effect on the stock composition of the catches.

Generally, stock proportions during the fishing season in each fishery area were more similar among years within each period (June and July-August) rather than between periods within years most likely reflecting migration patterns of different salmon populations. Catch compositions during the first period

(June) were more diverse and consisted of various populations from wide geographical area, whereas during the second period (July-August) local salmon populations contributed to the fishery at larger extent.

A relatively stable stock composition across the years within each period was observed in the majority of the studied fishery areas (Sør-Varanger; Nesseby and Vadsø; Nordkapp, Lebesby and Porsanger; Loppa and Hasvik; and Alta). As exception, the composition of Sør-Varanger catches in 2021 was presented by larger proportion of eastern Varanger P. populations during the last weeks of the fishing season compared to the previous years, which was most likely due to the absence of samples from the easternmost fishery locations of the Sør-Varanger in 2021. In addition, contribution of salmon populations from the northern Kola Peninsula, including Kola Bay, to the Sør-Varanger fishery has been declined compared to previous years. The most stable catch composition was observed in inner Alta fjord fishery area, where salmon from the R. Alta was dominating (77.8% - 93.5%) during the fishing season and across the years.

We observed a drastic decline of Tana salmon contribution to the fishery in Tana fjord and in adjacent municipalities (Vardø, Båtsfjord, Berlevåg and Gamvik). Thus, in 2008 – 2012 stock composition in Vardø, Båtsfjord, Berlevåg and Gamvik fishery area was mainly presented by relatively even contribution of salmon from northern Varanger P. rivers and populations of R. Tana system. However, in 2020 the proportion of Tana populations in catches was very low, decreased nearly 10 times, from 47.4% in 2008 to 4.7% in 2020. Similarly, in 2020 the overall proportion of Tana origin fish in Tana fjord catches was ca. 60%, which was much lower compared to the previous years (ca. 80-90%). A declining contribution of R. Tana populations to the catches observed in Tana fjord and adjacent areas was in line with observations of decreasing number of ascending adults in Tana fjord and R. Tana system during the recent years.

To get further insights into the status and dynamics of Atlantic salmon populations in the area, continuous catch composition monitoring of Norwegian commercial fishery in northern Norway and of scientific fishery in the Russian waters of the Barents Sea, as well as monitoring of juvenile densities in the rivers of northern Norway and the Kola P., is strongly recommended.

Keywords:

Salmo salar, microsatellites, baseline, genetic stock identification, catches, GSI

Table of Contents

Introduction	5
Material and Methods	6
<i>Baseline data</i>	6
<i>Adult Atlantic salmon data</i>	7
<i>DNA extraction and genotyping</i>	7
<i>Statistical analyses of the baseline samples</i>	8
<i>Baseline evaluation for GSI</i>	9
<i>Genetic stock identification</i>	10
Results	13
<i>Microsatellite loci in genetic baseline</i>	13
<i>Temporal stability among baseline samples collected in different years</i>	14
<i>Genetic variation of the baseline samples</i>	15
<i>Population genetic structure</i>	16
<i>Updated baseline GSI accuracy</i>	17
<i>GSI accuracy for populations of interest</i>	17
<i>Accuracy of stock contribution estimates for populations of interest</i>	18
Genetic stock identification of fishery samples	19
<i>Sør-Varanger fishery area</i>	19
<i>Nesseby and Vadsø fishery area</i>	26
<i>Vardø, Båtsfjord, Berlevåg and Gamvik fishery area</i>	27
<i>Tana fishery area</i>	30
<i>Nordkapp, Lebesby and Porsanger fishery area</i>	33
<i>Loppa-Hasvik fishery area</i>	33
<i>Alta fishery area</i>	33
Discussion	38
<i>Genetic structure and temporal variation of the updated baseline</i>	38
<i>GSI power of the updated baseline</i>	38
<i>GSI of fishery samples</i>	39
Conclusions	41
Acknowledgements	41
Supplementary data	42
References	43

Introduction

Atlantic salmon has always been an important exploitable resource for people living both along the rivers and the coast of northern Europe. Fisheries for salmon, while originally conducted for sustenance, have in the last century become a popular pastime for people from Norway and many other countries. This recreational fishery has provided new ways of generating income and livelihoods for people living in these areas, while salmon remains important as a food source (Lankia et al. 2022).

Managing Atlantic salmon in a sustainable way entails preserving the many populations in the northern areas for the future while at the same time ensuring that salmon remains available both for commercial and recreational fisheries to the benefit of both local people and visitors. Considering the various fishing methods in use, and the still limited knowledge on how different populations are exploited in the fisheries, this is a challenging task. Adding to the complexity is geographical and temporal variations in marine survival and growth of salmon populations, and we are now only beginning to understand the underlying factors that are driving the variations.

Fishery is one of the major factors affecting the production and survival of salmon populations. Moreover, coastal and riverine fisheries exploit salmon along its migration routes and the fisheries are considered to be mixed-stock fisheries, i.e. salmon from different populations may contribute to the catches. This is certainly the case for coastal fisheries that may intercept the migration routes of salmon from many rivers, while in-river fisheries can be considered to be mixed-stock fisheries in rivers that contain several distinct spawning entities/sub-populations, such as the Tana river. Mixed-stock fisheries may represent a risk to biodiversity and may cause overexploitation of populations of low abundance if the stock composition of mixed harvests is unknown (Crozier et al. 2004). Thus, establishing the origin of fish and time of migration of different populations is an important step towards sustainable fishery management. Application of genetic stock identification (GSI) methods allows for obtaining the knowledge of the level of contribution of different salmon populations to the catches, which in turn enables to infer exploitation rates of various stocks by the fishery. During the recent years GSI assisted in revealing Atlantic salmon stock composition in North American populations (Bradbury et al. 2015, Bradbury et al. 2016a), the mixed-stock fishery at Greenland (Bradbury et al. 2016b), as well as in historical catches around the Faroe Islands (Gilbey et al. 2017), in development of stock-specific coastal migration models for the four largest salmon populations (RR. Målselva, Alta, Tana and Kola) in the Barents and Norwegian Seas (Svenning et al. 2019) and in discovering of sea-age group relationships with migration patterns (O'Sullivan et al. 2022).

Both the current CoASal (KO4178), and the previous Kolarctic Salmon projects (KO197), aimed at reducing the uncertainties in some of the factors that present a challenge to managers by investigating the coastal migration and exploitation of both Russian and Norwegian salmon stocks in the northern areas using GSI. To define the origin of fish caught at sea the GSI approach compares their genetic profiles with genetic profiles of each population in the baseline. However, GSI accuracy and precision largely depend on how well genetic profiles of populations are represented in the baseline. In the previous Kolarctic Salmon project, a comprehensive genetic baseline for Atlantic salmon populations from the R. Beiarelvva in Nordland to R. Pechora in Russia was generated (Ozerov et al. 2017). A wide-ranging sampling program was undertaken, often conducted under challenging conditions in remote areas was undertaken to establish as comprehensive and representative database of the genetic variation in these northern stocks. Building on this, the current CoASal project has expanded and updated this genetic baseline by revisiting a selection of the rivers in the baseline, and new samples were collected and analyzed. This work was undertaken to evaluate whether the genetic profiles of salmon populations are stable over time, and also to increase the precision of GSI.

In this project, we aimed to investigate whether exploitation patterns have changed since the previous investigation period, with special attention on Varanger fjord area (Sør-Varanger, Nesseby and Vadsø municipalities). This area can be considered as a reference area, as it was shown that interceptory fishery here has been the most intensive both for salmon originating from the rivers of Finnmark in Norway and from the Kola Peninsula in Russia (Vähä et al. 2014). In addition, we

compared recent and previous catch compositions in other Finnmark municipalities (spanning from Vardø to Alta municipalities). Further, as the management regime for coastal fisheries was adjusted in the region based on the findings of Kolarctic Salmon project, we aimed to explore if this have had an effect on the stock composition of the catches.

Material and Methods

Baseline data

To supplement the baseline generated during the Kolarctic salmon project (KO197), which included more than 180 Atlantic salmon populations from the rivers spanning from the Russian Archangelsk region in the East to the Norwegian Norland county in the West (Ozerov et al. 2017), we sampled additional rivers/year-classes with the special attention to the rivers in cross-border area. In total, 3082 Atlantic salmon parr (0–4 year age) were collected by electrofishing from 27 rivers/river locations in Northern Norway during 2019 and 2020 (Fig. 1; Table S1). We followed the stratified random sampling procedure to obtain a representative baseline population sample as described earlier (Ozerov et al. 2017). In brief, for every river, 2–6 separate areas or sections within the watershed, that were not adjacent to each other, were sampled, usually extending over a 100 m stretch. If possible, juveniles from multiple year classes were sampled from each area. All the juveniles collected were transferred to containers with water (buckets), where they were kept alive in cool aerated water for a short period. Fish that was sampled, were swiftly killed with adequate blow to the head, before putting on ice. Further, each juvenile was placed in an individual zip-locked plastic bag and immediately snap frozen on dry ice (see the KO4178 CoASal Sampling manual for details; Ozerov et al. 2020). Later the frozen samples were transported to the laboratory for further processing. A tissue sample from 2376 individuals were collected and stored in ethanol for later genetic analyses. The permits required to obtain samples were issued by the County Governor of Troms and Finnmark and Norwegian Environment Agency (Norway). In addition, we included in the baseline 413 individuals from six rivers (Repparfjordelva, Kvalsundelva, Lakselva Kviby, Transfarelva, Alta-Eibyelv and Halselva) collected in 2017 and genotyped using the same set of 31 microsatellites at the Institute of Marine Research (IMR, Norway; Project 15696-07).

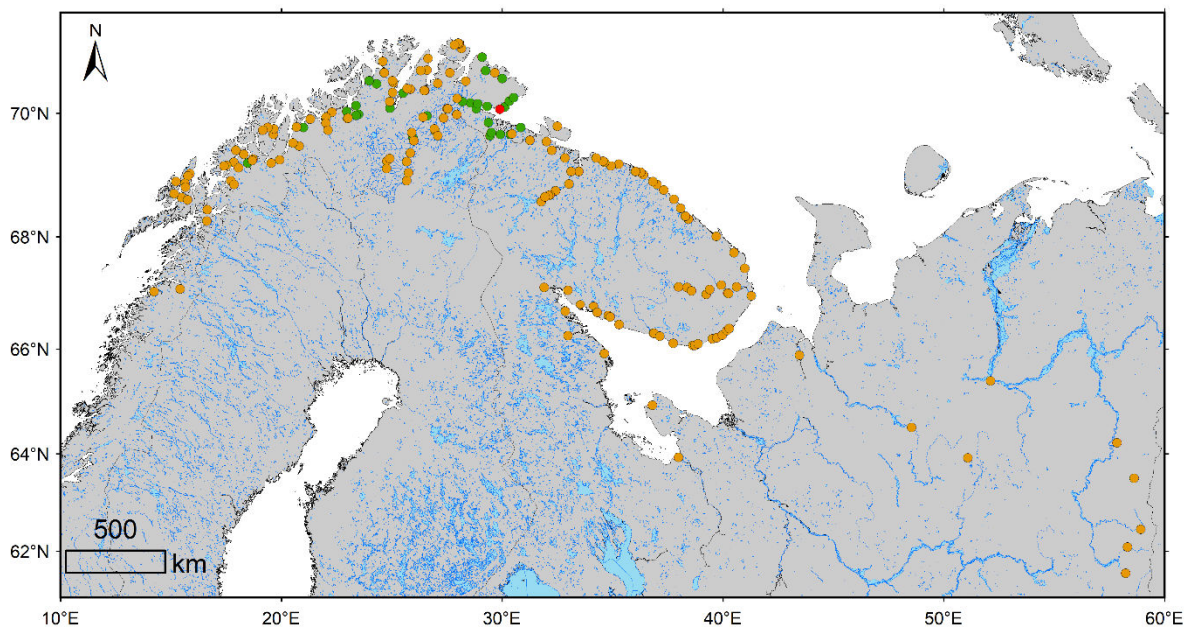


Fig. 1. Map indicating sampling locations. Green, orange and red circles represent the updated baseline samples, baseline samples collected earlier and Sjøbuselva sample excluded from the analyses due to high number of full-siblings, respectively.

Adult Atlantic salmon data

A total of 3258 adult wild Atlantic salmon were sampled in 2020 and 2021 (Table 1) along ~550 km of the North-Norwegian coast from 21.6 to 30.8°E and 69.7 to 70.8°N (Fig. 2). Samples were obtained from local professional fishers using commercial fishing gear (bend nets and/or bag nets) during the ordinary fishing season.

Table 1. Number of wild Atlantic salmon captured monthly in the ordinary coastal fishery season in Northern Norway in 2020 and 2021, and the numbers of Atlantic salmon genotyped and used for GSI.

Year	Wild adult salmon	June	July	Aug	Total
2020	Number of salmon captured	1368	518	n/a	1886
	Number of salmon genotyped	1351	516	n/a	1867
	Number of salmon GSI	1259	484	n/a	1743
2021	Number of salmon captured	800	547	7	1354
	Number of salmon genotyped	790	543	7	1340
	Number of salmon GSI	623	487	4	1114

For each individual location, capture method, and date were recorded. Body mass (g) and total length (cm) were measured, sex was determined and a scale sample was taken for: (i) age and growth determination, (ii) genetic analysis, and (iii) estimation of the number of escaped farmed fish in the catch. Based on scale pattern analysis fish were categorized as wild or escaped farmed fish, and only wild fish were subsequently used for genetic analysis.

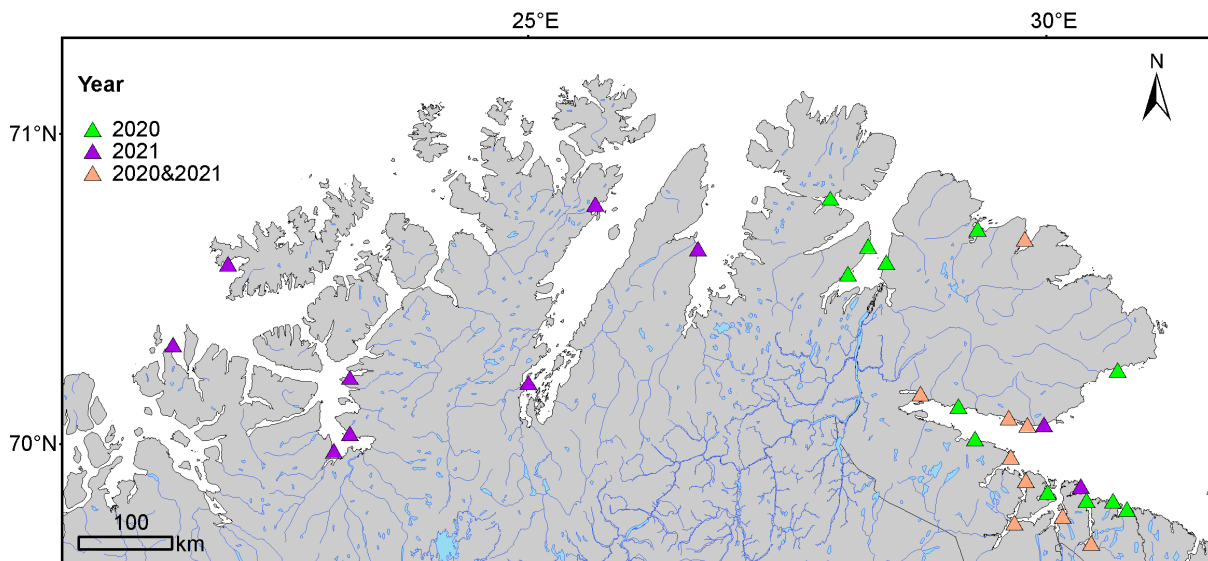


Fig. 2. Map showing the fishing locations in 2020 and 2021. Green, violet and cantaloupe filled triangles represent fishing locations in 2020, 2021 and overlapped locations in 2020 and 2021, respectively.

DNA extraction and genotyping

Total genomic DNA was extracted at the University of Turku (UTU, Finland) from juvenile fin tissue or adult scale according to (Elphinstone et al. 2003) or at the Institute of Marine Research (IMR, Norway) from juvenile fin tissue using Qiagen DNeasy 96 Blood & Tissue kits (Qiagen™), following the manufacturer's recommendations. Each sample was surveyed for genetic variation at 31 microsatellite DNA loci (see Ozerov et al. 2017 for details). To minimize genotyping errors, electropherograms and allele scores were reviewed by two persons independently. The genotyping

quality threshold was initially set to having 27 of 31 loci producing unambiguous data with failure resulting in re-analysis from either the DNA extraction or PCR amplification step.

Given that microsatellite genotype data were generated at two different laboratories (UTU and IMR), we performed a calibration and standardization of allele calling. Calibration included both laboratories genotyping the same 344 samples from 10 baseline populations (Table 2), comparing genotypes and standardizing allele bins as well as exchanging pictures of amplicon profiles from each locus and the majority of alleles to standardize allele calling.

Table 2. Populations used for calibration of microsatellite loci, their geographical locations and number of individuals (*n*).

<i>Population</i>	<i>Year</i>	<i>Location</i>	<i>n</i>
<i>Norway</i>			
Grense Jakobselv	2019	69.73°N 30.89°E	50
Klokkarelv	2019	69.86 °N 29.39°E	24
Vesterelva	2019	70.12 °N 28.51°E	36
Vestre Jakobselv	2019	70.14 °N 29.32°E	35
Skallelva	2019	70.18 °N 30.25°E	20
Komagelva	2019	70.24 °N 30.51°E	29
Syltefjordelva	2019	70.52 °N 29.42°E	19
Kongsfjordelva	2019	70.59 °N 29.09°E	79
Storelva/Berlevåg	2019	70.84 °N 29.04°E	31
Sjøbuselv	2019	70.06 °N 29.90°E	21
Total			344

In total, 234 individuals in the baseline were discarded due to low quality genotype score ($n = 175$), contamination ($n = 5$) or being potential triploids ($n = 2$). In addition, 61 individuals showed amplification of trout allele at SsaD486 locus being trout or salmon-trout hybrids and thus were excluded from further analysis.

Among the samples of adult individuals 401 samples were excluded from the further analyses due to low genotype score ($n = 328$), contamination ($n = 29$), being farmed escapee ($n = 14$), salmon-trout hybrid ($n = 14$), potential triploid ($n = 2$), potential pink salmon ($n = 6$) or genetic duplicate ($n = 8$).

Statistical analyses of the baseline samples

The sibship-reconstruction method implemented in Colony 2.0.6.6 (Jones and Wang 2010) was applied to test for full- and half-sib relationships in each location, as samples dominated by a few full-sib families may lead to biased estimates of allele frequencies in populations (Hansen et al. 1997), and may influence the results from Bayesian clustering algorithms employed for investigating population structure (Rodriguez-Ramilo and Wang 2012). To minimize the effect of family structure on estimation of genetic structure and genetic stock identification (GSI) all full-sibs except one pair per family were removed from the baseline samples, so that no family contained more than two full siblings. This resulted in exclusion of 152 individuals (5.97%). As the majority of the samples from the R. Sjøbuselv ($n = 35$) were represented by a large full-sib family ($n = 27$), this population was excluded from the further analyses. Finally, genetic profiles of 2,384 newly analyzed samples were joined with the previously generated baseline data of 12,860 specimens resulting in 15,244 individuals included in the updated baseline for further analyzes.

Given that some populations were repeatedly sampled in different years (Table 3), the stability of the population structure was examined by comparing the temporal variation within rivers with the spatial variation among rivers by applying a two-level hierarchical analysis of molecular variance (AMOVA) with 10,000 permutations using Arlequin 3.5.2 (Excoffier and Lischer 2010). Given that temporal variation (0.58%) was ca. 6.5 times lower compared to spatial variation (3.78%, see the

Results, Table 6), temporal replicates at each river/river location were combined for further analyses (see Table 3).

The presence of potential genotyping errors and null alleles was tested using *null.all* function of PopGenReport 3.0.4 package (Adamack and Gruber 2014) in R 4.0.5 (R Core Team 2021). Basic genetic diversity of microsatellite loci and populations (observed and unbiased expected heterozygosity and allelic richness) were calculated using the R-package *diveRsity* 1.9.90 (Keenan et al. 2013). The same package was used to estimate Weir and Cockerham (1984) statistics, and to test for deviations from Hardy–Weinberg equilibrium (HWE) for every locus-sample combination using 999,999 Monte Carlo replicates. Statistical significance levels for HWE were corrected for false discovery rate (FDR; Benjamini and Hochberg 1995) implemented in the *p.adjust* function in the *stats* package of R 4.0.5. Pairwise F_{ST} -values among the populations and their significance was estimated using 1000 permutation test in *GenoDive* 3.0 (Meirmans and Van Tienderen 2004). Average pairwise F_{ST} among 20 neighbouring populations was used as a proxy for population divergence in the region, i.e. population-specific F_{ST} . The effective population sizes (N_e) were estimated using the linkage disequilibrium method (Hill 1981) as implemented in *NeEstimator* 2.1 (Do et al. 2014) applying the 0.05 exclusion criteria for allele frequencies (Waples 2006).

The genetic structure of the updated baseline populations was estimated using Bayesian clustering as implemented in *Structure* 2.3.4 (Pritchard et al. 2000, Falush et al. 2003) applying parallelization at CSC HPC cluster (IT Center for Science Ltd., Finland) using the same parameters as described in Ozerov et al. (2017). In brief, a hierarchical approach was applied (e.g., Rosenberg et al. 2002, Vähä et al. 2007) where the dataset was explored sequentially, by first identifying major genetic shifts/clusters, dividing the dataset according to the identified major groups, and then repeating the analysis independently on these smaller datasets. We repeated the analyses of the whole dataset and after definition of the major genetic clusters, which remained the same as in Ozerov et al. (2017; see Results), we further explored fine-scale genetic structure only in the major genetic clusters supplemented with new genetic data (3 - Northern Kola Peninsula/Southern Varanger fjord; 4 - Varanger Peninsula; 5 - R. Teno system; 6 - Western Finnmark and 7 - Troms-Nordland). For each round of the *Structure* analysis, the algorithm was run 10 times for each K to ensure convergence of values and with 750,000 MCMC repeats preceded by a burn-in of 250,000 steps for each K = 1 to 8. The datasets were analyzed using an admixture model with correlated allele frequencies. *Structure* outputs were further compiled in *Structure Harvester* 0.6.94 (Earl and vonHoldt 2012), before replicate runs were combined in the *Clumpp* 1.1.2 (Jakobsson and Rosenberg 2007). Microsoft Excel was used to construct plots of the merged replicates of population clustering from *Clumpp*.

Given the observed temporal stability of genetic structure and the same large- and fine-scale genetic clustering revealed with *Structure* (see Results), we kept the same reporting groups (i.e. the group of populations for GSI purposes) as in Ozerov et al. (2017) for further analyses. The large-scale regional RG_R included: 1 - Eastern Barents Sea; 2 - White Sea; 3 - Northern Kola Peninsula/Southern Varanger fjord; 4 - Varanger Peninsula; 5 - R. Teno system; 6 - Western Finnmark; 7 - Troms-Nordland. Regional RGs were further subdivided into 26 local RG_L on the basis of genetic clustering and geographical location (Table S1).

Baseline evaluation for GSI

The performance of the updated baseline for accurate GSI was assessed using the previously described pipeline (Ozerov et al. 2017). In brief, a conditional maximum likelihood (CML) based approach as implemented in *ONCOR* (Kalinowski et al. 2008) and the Bayesian methodology of Pella and Masuda (2001) implemented in *cBayes* 5.0.1 (Neaves et al. 2005) were used. For the latter approach, six independent 100,000 iteration Monte Carlo Markov chains were produced. To remove the influence of initial starting values, only the last 10,000 iterations of each chain were combined and used to estimate individual assignment to population and RG.

First, 100% simulation method (Anderson et al. 2008) implemented in ONCOR (Anderson et al. 2008) was applied to test the ability of the updated baseline to accurately assign fish to their populations and RGs of origin. The mixture sample sizes were set to 200 and simulations were repeated 100 times for both the individual populations as well as for the defined RGs.

Second, to further evaluate the baseline performance, hypothetical fishery samples were generated by random sampling of 1000 individuals of known origin from the baseline (Ackerman et al. 2011). The hold-out mixtures were analyzed using both CML and Bayesian approaches. The analyses were repeated 20 times with different hold-out mixtures.

Third, to determine GSI success rate in a mixture, more detailed simulations were performed for 20 populations, which were selected due to their largest stock sizes in the study area (SSC = 4–5, Table S1) and considered as the most important contributors to the coastal fisheries in the northern Norway. The mixtures were generated by random removal of 10 baseline individuals of known origin from each of 20 populations of interest ($n = 200$) and joined with one of four mixtures consisting of 679–2317 individuals simulated using ONCOR. Thus, each mixture sample consisted of 200 resampled individuals from the 20 populations of interest and simulated individuals from the populations in the region. Both the simulated and resampled mixtures were randomly replicated 40 times. Further, the success of correct assignments for the 20 populations of interest was evaluated in two different scenarios. The first, a regional fishery scenario, when the contribution of populations of interest to the mixture was relatively high (4–24%), putatively reflecting the composition of fisheries catch from inner coastal areas during the time close to spawning season. Another is non-regional fishery scenario, the contribution of populations of interest to the mixture was relatively low (0.5–1.5%), putatively reflecting the composition of fisheries catch in outer coastal areas.

Genetic stock identification

Population and RG of origin for each successfully genotyped adult individual was estimated using the Bayesian GSI methodology described in Pella and Masuda (2001) and implemented in cBayes 5.0.1 (Neaves et al. 2005) using six independent chains of 100,000 iterations. The last 10,000 iterations of each chain were combined and used to estimate individual assignment to the population of origin to remove the influence of initial starting values.

The population composition of the mixture sample was expected to influence the estimates of stock composition as the method utilizes this information during assignment (Pella and Masuda 2001). Thus, to improve the sensitivity of stock estimates (Vähä et al. 2017) large mixture samples of each year were subdivided into subsets based on location and time of catch. Given that the number of samples of the 47 localities per month in 2020 and 2021 were small, they were grouped into 14 analysis regions and two time periods, covering official fishing season: period 1 (June) and period 2 (July – August; Table 3).

To compare catch compositions of 2020 and 2021 with those observed during previous years in Finnmark county, the adult salmon samples caught in 2008, 2009, 2011 and 2012 coastal fisheries (see Vähä et al. 2014, Svenning et al. 2019) were grouped into subsets using the approach described above and re-analyzed using the updated baseline. However, the period of the ordinary fishing season differs in 2008-2009 and 2011-2012 and some fishers were allowed to fish beyond the official fishing season in 2008-2009 and 2011-2012, i.e. from early May until early September in these 4 years. Thus, the previous coastal samples caught during the periods corresponding to 2020 and 2021 actual fishing times in Finnmark municipalities were grouped into two time periods: period 1 (June) and period 2 (July – August) and the samples beyond these periods were grouped into period 0 (May) and period 3 (mid August – September) and analyzed separately (Table 3). In addition, the samples caught in Troms and Norland 2011-2012 coastal fisheries were re-analyzed using the updated baseline with periods 1 and 2 corresponding to 2011-2012 ordinary fishing times (weeks 21-32), whereas the samples caught in earlier May (weeks 18-20) and mid-August – earlier September (weeks 33-37) were grouped into periods 0 and 3, respectively. Collectively, the 23,615 samples were divided into 238 temporally and spatially distinct subsets for analysis, including 35 subsets, which were combined within each year or

periods within a year (Table 3). The probability (p) threshold for assignment of an individual to a population or RG was ≥ 0.7 and was applied at population and RGs levels as recommended earlier (Vähä et al. 2011, Vähä et al. 2014, Bradbury et al. 2015). Further, we present only the results for Finnmark fishery during the official fishing time (periods 1 and 2). Given small sample size in some areas and periods after applying the probability threshold, the results of GSI were combined into seven fishery areas overlapped across all analyzed years: 1) Sør-Varanger; 2) Nesseby and Vadsø; 3) Vardø, Båtsfjord, Berlevåg and Gamvik; 4) Tana; 5) Nordkapp inner, Lebesby and Porsanger; 6) Loppa and Hasvik; 7) Alta. The updated genetic assignments of samples caught beyond the official fishing season (period 0 and 3) and samples caught in other areas of Finnmark, as well as in Troms and Nordland can be found in Supplementary file 1. All the pie and bar plots were generated using ggplot2 (Wickham 2016) library in R. The maps were generated in ArcMap 10.5.1 (ESRI 2015).

Table 3. Coastal fishery samples arranged in 238 subsets for GSI analyses based on their spatial and temporal distribution.

Region name	Year																				Total
	2008				2009				2011				2012				2020		2021		
	Period																				
	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	1	2	1	2	
01.Sør-Varanger-East	36	152	47						145	150	96		70	337	117		208	24	142		1524
02.Sør-Varanger-West	210	512	264	26	114	34			186	471	556	11	79	469	587	64	428	57	167	104	4339
03.Nesseby-Fjord	59	86	67	27	130	31			83	68	100	46	7*	152*	226*	8*	387	199	46	135	1857
04.Vadsø									91	39	180	67	66	114	252	22	52	23	61 [†]	2 [†]	969
05.Vardø	10	36	30*	3*			10						27	20	64	17	8 [†]	26			251
06.Båtsfjord-Berlevåg	20	96	34		21	15			8*	42*	73	18	1*	104*	63	23	123	103	1		745
07.Tana	140	200	119		43	26			26	64	51		25	53	179		45	50			1021
08.Gamvik	33	16	18	30	19	16	19*	2*	40	40*	6*		1 [†]	5 [†]	3 [†]	9 [†]	7 [†]	2 [†]			266
09.Lebesby	118	210	151	23					69	62	214	39	36	93	279	260			15	61	1630
10.Nordkapp-Inner	8*	15*	65	54	29	40	29		32	136	97	132	27	89	201	224			28	15	1221
11.Nordkapp-Outer	54	94	34	37					46	54	53	32	63	170	103	89					829
12.Porsanger									8 [†]	7 [†]	4 [†]	4 [†]	4*	22*	133	194			25	17	418
13.Måsøy	6 [†]	24 [†]	1 [†]						1*	8*	27	19	80	67	44	78					355
14.Kvalsund-Hammerfest									28	53	69*	4*	13	25	32	27					251
15.Hasvik	89	110			10 [†]	3 [†]	3 [†]		108		41*	4*	32	91	13	17			84	16	621
16.Loppa	73	103	234	58	19	83*	1*		55	28	126		75	64	206	26			25	50	1226
17.Alta			38		1 [†]	95 [†]			9*	35*	251*	3*	2*	121*	183	19			29	91	877
18.North-Troms-North inner									5*	182*	282*	5*	1*	235*	384	36					1130
19.North-Troms-North outer													1*	85*	138*	6*					230
20.South-Troms-Målselva									2*	311*	458	23	4	476	557	67					1898
21.South-Troms-Middle outer									31	250	166		6	156	155						764
22.South-Troms-North outer									1*	78*	86			158	113						436
23.Nordland-North										59	61*	2*	10	98	37*	7*					274
24.Nordland-South									3*	34*	32*	3*	10	274	114	13					483
Total	856	1654	1102	258	20	382	356	35	977	2171	3029	412	640	3478	4183	1206	1258	484	623	491	23615

Samples with low number of individuals were combined within a year ([†]) or larger periods within a year (*) and region.

Results

Microsatellite loci in genetic baseline

Deviations from Hardy–Weinberg equilibrium ($p < 0.01$) were detected at 154 out of 5635 locus-population combinations. After correction for FDR, six of the combinations remained significant at the 1% level (Table S2). The potential presence of null-alleles was detected at 88 locus-sample combinations out of 5766 tests (Table S3). However, the presence of loci prone to null-alleles most likely does not alter the overall outcome of GSI, although such loci may lower assignment power and result in a slight overestimation of genetic divergence (Carlsson 2008). Regardless, as none of the loci showed consistent HWE deviations or potential presence of null alleles across multiple populations, we retained all 31 microsatellites for further analysis.

Table 4. Overall microsatellite diversity in baseline samples. Observed (H_O) and expected (H_E) heterozygosity, allelic richness (A_R), number of alleles and genetic divergence indices with 95% confidence intervals: F_{ST} (Weir and Cockerham 1984), G'_{ST} (Hedrick 2005) and D (Jost 2008).

Locus	H_O	H_E	A_R	N_{alleles}	F_{ST} (95%CI)	G'_{ST} (95%CI)	D (95%CI)
<i>EST28</i>	0.332	0.333	3.08	15	0.046 (0.042-0.051)	0.078 (0.070-0.086)	0.027 (0.024-0.031)
<i>EST68</i>	0.611	0.619	4.48	8	0.055 (0.051-0.058)	0.174 (0.163-0.185)	0.115 (0.108-0.123)
<i>MHCII</i>	0.710	0.720	6.21	18	0.085 (0.082-0.088)	0.330 (0.320-0.340)	0.262 (0.253-0.270)
<i>SSA405</i>	0.920	0.919	14.69	33	0.028 (0.027-0.030)	0.409 (0.397-0.421)	0.389 (0.377-0.400)
<i>SSf43</i>	0.393	0.393	3.64	12	0.043 (0.039-0.046)	0.070 (0.063-0.077)	0.029 (0.025-0.032)
<i>SSsp2215</i>	0.887	0.890	11.85	31	0.034 (0.033-0.036)	0.346 (0.335-0.356)	0.320 (0.310-0.330)
<i>SSsp2216</i>	0.876	0.872	10.50	24	0.030 (0.029-0.032)	0.274 (0.261-0.287)	0.248 (0.236-0.259)
<i>SSsp3016</i>	0.796	0.789	9.18	22	0.044 (0.041-0.046)	0.232 (0.221-0.244)	0.193 (0.184-0.203)
<i>Ssa14</i>	0.457	0.453	2.14	7	0.093 (0.087-0.100)	0.173 (0.161-0.186)	0.087 (0.081-0.093)
<i>Ssa171</i>	0.812	0.812	8.69	29	0.045 (0.043-0.047)	0.273 (0.262-0.284)	0.234 (0.224-0.244)
<i>Ssa197</i>	0.889	0.890	12.12	31	0.038 (0.036-0.039)	0.397 (0.386-0.407)	0.369 (0.359-0.379)
<i>Ssa202</i>	0.827	0.824	8.49	20	0.055 (0.052-0.057)	0.336 (0.325-0.347)	0.294 (0.285-0.304)
<i>Ssa289KA</i>	0.635	0.645	3.91	12	0.075 (0.071-0.079)	0.232 (0.220-0.243)	0.163 (0.155-0.171)
<i>Ssa412</i>	0.547	0.545	3.18	10	0.117 (0.113-0.122)	0.288 (0.277-0.299)	0.181 (0.174-0.188)
<i>Ssa98</i>	0.396	0.397	4.03	17	0.047 (0.043-0.050)	0.092 (0.084-0.100)	0.039 (0.035-0.043)
<i>SsaD486</i>	0.012	0.012	1.16	6	0.022 (0.013-0.032)	0.026 (0.014-0.041)	0.000 (0.000-0.001)
<i>Ssosl25</i>	0.751	0.746	6.28	20	0.054 (0.052-0.057)	0.244 (0.233-0.254)	0.194 (0.185-0.203)
<i>EST107</i>	0.641	0.638	4.40	10	0.050 (0.046-0.053)	0.133 (0.123-0.143)	0.089 (0.082-0.096)
<i>EST19</i>	0.878	0.881	12.28	43	0.038 (0.036-0.040)	0.353 (0.342-0.365)	0.325 (0.314-0.337)
<i>MHC I</i>	0.796	0.803	8.14	21	0.058 (0.055-0.061)	0.295 (0.284-0.306)	0.252 (0.242-0.262)
<i>SSsp1605</i>	0.752	0.748	6.13	13	0.055 (0.053-0.058)	0.246 (0.234-0.258)	0.197 (0.187-0.206)
<i>SSsp2201</i>	0.920	0.921	12.42	40	0.025 (0.024-0.026)	0.368 (0.357-0.381)	0.350 (0.338-0.361)
<i>SSsp2210</i>	0.783	0.782	7.41	17	0.062 (0.060-0.065)	0.313 (0.302-0.324)	0.263 (0.254-0.272)
<i>SSspG7</i>	0.842	0.840	10.05	25	0.038 (0.036-0.040)	0.283 (0.270-0.296)	0.249 (0.237-0.261)
<i>Sleel53</i>	0.499	0.505	3.41	11	0.078 (0.073-0.083)	0.186 (0.174-0.198)	0.104 (0.097-0.111)
<i>Sleen82</i>	0.730	0.735	5.58	12	0.048 (0.046-0.051)	0.214 (0.203-0.226)	0.167 (0.158-0.176)
<i>Ssa407</i>	0.872	0.878	11.42	41	0.028 (0.026-0.029)	0.245 (0.234-0.256)	0.222 (0.212-0.232)
<i>SsaD144</i>	0.925	0.927	15.21	39	0.027 (0.026-0.029)	0.431 (0.420-0.443)	0.413 (0.402-0.424)
<i>SsaD157</i>	0.913	0.911	6.05	42	0.022 (0.021-0.023)	0.312 (0.297-0.328)	0.292 (0.279-0.307)
<i>Ssleer15.1</i>	0.618	0.625	1.62	7	0.079 (0.075-0.083)	0.249 (0.237-0.261)	0.171 (0.163-0.180)
<i>Ssosl85</i>	0.793	0.790	7.85	26	0.053 (0.051-0.055)	0.269 (0.258-0.281)	0.225 (0.216-0.235)
<i>Total</i>	0.704	0.705	7.15	662	0.050 (0.049-0.051)	0.193 (0.190-0.196)	0.171 (0.169-0.173)

In total, 662 microsatellite alleles were observed among baseline samples, ranging from six (*SsaD486*) to 43 (*EST19*) alleles per locus (mean = 21; Table 4). The expected heterozygosity (H_E) across the baseline samples varied from 0.012 (*SsaD486*) to 0.927 (*SsaD144*) with across loci average of 0.705. The mean allelic richness (A_R) of loci was 7.15 and varied from 1.16 (*SsaD486*) to 15.21

(SsaD144). The mean genetic divergence across all loci and samples estimated as F_{ST} was 0.050, varying from 0.022 (SsaD486) to 0.117 (Ssa412; Table 4).

Temporal stability among baseline samples collected in different years

In some rivers and tributaries, samples were collected in different years. In total 42 rivers/tributaries were sampled at different time points, with a time span between samplings from one to 12 years. Genetic differentiation (F_{ST}) among temporal samples was significant ($P < 0.01$) in 16 of the 42 rivers/tributaries (Table 5).

Table 5. Genetic differentiation (F_{ST}) between temporal samples in the baseline and corresponding P -values.

<i>Rivername</i>	<i>Sampling year</i>	<i>F_{ST}</i>	<i>P</i>
Pechora-Unya	2000/2012	0.020	0.000
Verkhnaya Pechora	2001/2012	0.010	0.000
Pechora-Ilych	2002/2003/2007	-0.008	0.953
Kuloy-Soyana	2006/2008	-0.002	0.700
Keret	2011/2012	0.021	0.000
Ponoi main stem	2008/2009	0.000	0.517
Grense Jakobselv	2006/2010/2019	0.006	0.000
Karpelv	2009/2019	0.005	0.000
Sandneselva Kirkenes	2009/2019	0.007	0.000
Munkelv	2009/2019	-0.001	0.590
Neiden	2006/2010/2019	-0.013	1.000
Klokkarelv	2009/2019	0.002	0.053
Nyelva	2010/2019	0.002	0.212
Vesterelva	2010/2019	0.002	0.011
Bergebyelva	2009/2019	0.000	0.330
Vestre Jakobselv	2006/2010/2019	0.004	0.000
Storelva Vadsø	2010/2019	0.013	0.000
Skallelva	2009/2019	0.000	0.332
Komagelva	2006/2019	0.000	0.495
Syltefjordelva	2009/2019	0.009	0.000
Kongsfjordelva	2007/2019	0.021	0.000
Storelva/Berlevåg	2012/2019	-0.006	1.000
Teno-Iesjohka	2006/2007	0.003	0.000
Teno-Karasjohka	2006/2009	-0.029	1.000
Teno-Valjohka	1997/1999/2010	-0.028	1.000
Teno-Utsjoki	1998/1999	-0.001	0.182
Teno main stem, Outakoski	2010/2020	0.008	0.000
Teno main stem, Upper	2010/2020	-0.009	1.000
Teno main stem, Tana Bru	2010/2020	0.003	0.023
Børselv	2006/2010/2020	-0.012	1.000
Lakselva Porsanger	2006/2010/2020	0.013	0.000
Repparfjordelva	2006/2010/2017/2020	-0.001	0.295
Kvalsundelva	2010/2017	0.007	0.001
Lakselva Kviby	2010/2017	-0.013	1.000
Alta	2007/2010/2020	-0.001	0.389
Alta-Eibyelv	2010/2017	-0.009	1.000
Halselva	2011/2017	0.012	0.004
Reisa	2007/2011/2020	-0.001	0.846
Målselv	2007/2011/2020	-0.003	1.000
Roksdalsvassdraget	2007/2010	0.006	0.009
Alsvågvassdraget	2008/2011	0.003	0.115
Gårdselv	2007/2008	-0.004	0.839

However, F_{ST} values among temporal samples within rivers/tributaries were low (F_{ST} mean = 0.001, ranged from -0.029 to 0.021; F_{ST} median = 0.001) compared to the average among rivers (F_{ST} =

0.050). Moreover, the variation among samples due to temporal component (0.58%) was ~6.5 times lower than that due to the spatial component (3.78%; Table 6).

Table 6. Hierarchical analysis of molecular variance (AMOVA) in the temporal samples of the genetic baseline.

Source of variation	Sum of squares	Variance components	Percentage variation
Among spatial samples	4602.5	0.319	3.78
Among temporal samples	805.3	0.049	0.58
Among individuals within populations	51205.2	0.039	0.46
Within individuals	51505.0	8.035	95.17

Genetic variation of the baseline samples

The mean level of genetic diversity across loci varied considerably among baseline populations. Similar to the results observed earlier (Ozerov et al. 2017), genetic diversity was the highest in 063-Zarubikha Kildin ($H_E = 0.744$, $A_R = 7.61$) and 074-Titovka ($H_E = 0.734$, $A_R = 8.28$) on the northern coast of the Kola Peninsula (Barents Sea), and the lowest in 014-Kovda ($H_E = 0.552$, $A_R = 3.66$), Kandalaksha Bay of the White Sea (Table S1). In general, genetic diversity presented as allelic richness, increased along the coast from east to west, and was higher on the northern coast of the Kola Peninsula than in the White Sea. The level of genetic diversity slightly decreased among the populations of the Varanger Peninsula and was low among some of the R. Teno tributaries and the rivers of Gamvik area. Further, genetic diversity increased southward along the Norwegian coast (Fig. 3).

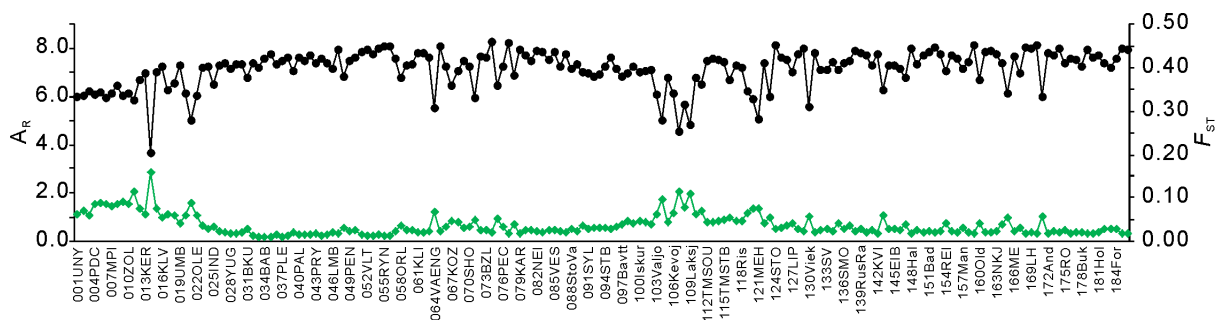


Fig. 3. Genetic diversity (estimated as allelic richness (A_R); black solid line) and genetic divergence (estimated as pairwise F_{ST} among 20 closest neighbouring populations; green solid line) across 186 Atlantic salmon populations.

In line with the pattern observed for genetic diversity, the highest genetic divergence was detected among the Eastern Barents and inner White Sea populations (001UNY–019UMB), and decreased further among the Kola Peninsula populations (020KUZ–063ZRK; Fig. 3). In the western part of the range, the highest genetic divergence was observed among the populations of the Kola Bay (064VAENG–071KLN), the River Teno system (104Galdd, 106Kevoj–110Levaj and 116Vetjo) and in the Gamvik area (119San–121MEH).

Estimates of N_e from the LD method varied from 8 (95%CI 8–9) in 160-Oldervikelva Troms to 4147 (95%CI 269–inf) in Ponoj tributary 039-Losinga with a mean = 221 and median = 103 individuals (Table S1). Pairwise genetic divergence (measured as F_{ST}) among 186 populations varied from 0.001 to 0.245 and was significant ($p < 0.01$) in 17 190 of 17 205 pairwise comparisons (Table S4). The highest pairwise F_{ST} value observed was 0.245, between 014-Kovda and 109-Laksjohka, and the lowest pairwise $F_{ST} = 0.001$ was observed within the R. Ponoj system, between 035-Ponoj mainstem and 044-Ponoj Tomba. Ten of the 15 non-significant comparisons were attributed to low genetic divergence mostly among the geographically close populations in the Troms-Nordland region ($F_{ST} = 0.003–0.006$), and within the river systems Ponoj ($F_{ST} = 0.001–0.005$) and Teno ($F_{ST} = 0.003$; Table S4).

Population genetic structure

The analyses of hierarchical genetic structure revealed distinct genetic clusters among the salmon populations with increasing number of K . As expected the genetic structure pattern was similar to this observed earlier (Ozerov et al. 2017) and at $K = 7$ the major genetic clusters largely corresponded to the large geographic regions (Fig. 4).

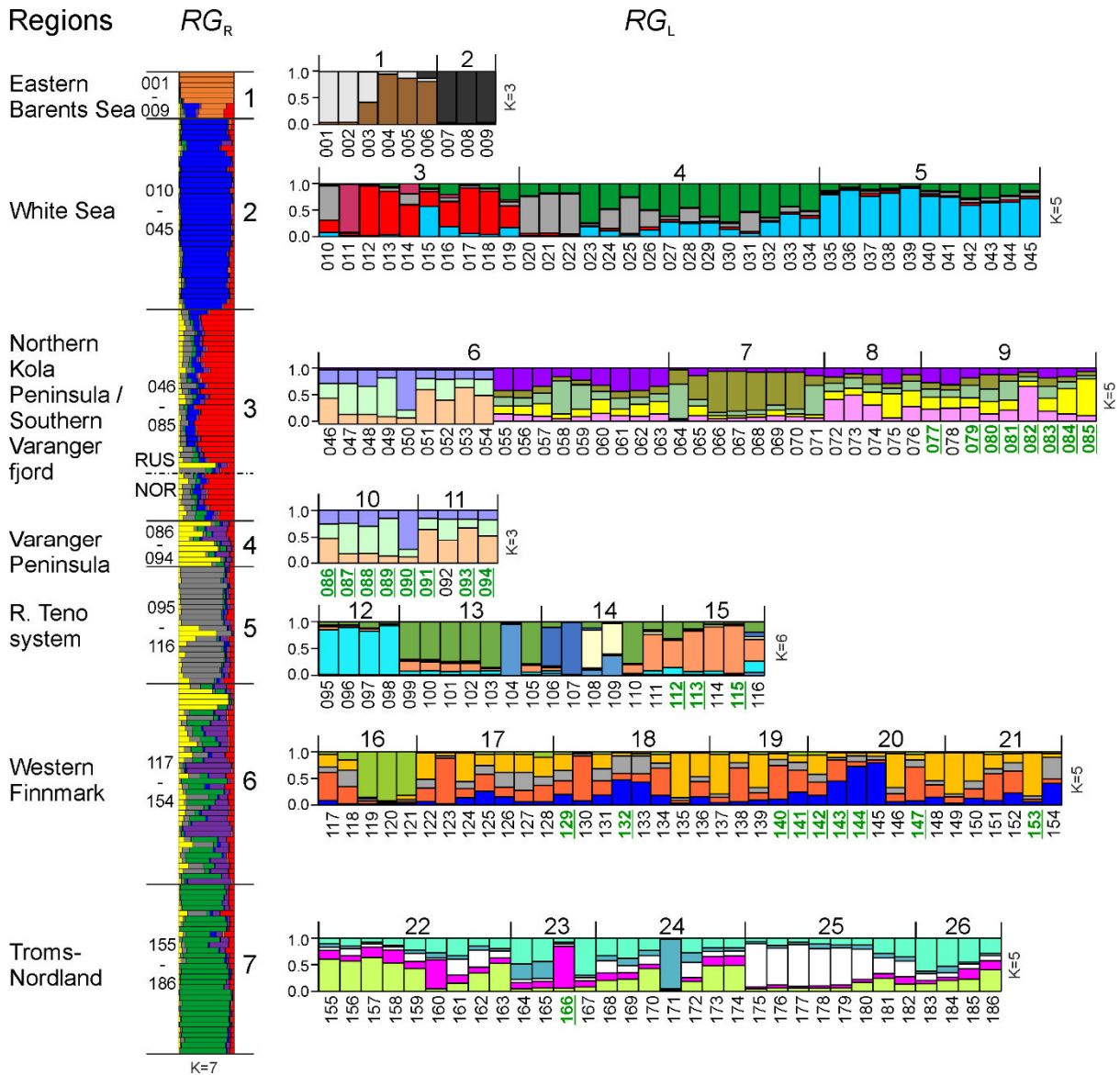


Fig. 4. Hierarchical Structure plot, indicating estimated major genetic structure shifts in the studied area. The numbers on the left/below the plots indicate population ID, the numbers on the right/above the plots indicate regional (RG_R) and local (RG_L) reporting groups. The IDs of populations supplemented with new genetic data are highlighted in underlined bold green. Note: populations 104 and 116 belong to the RG_L 14. Population IDs follow those given in Table S1.

Fine-scale genetic structure within the major genetic clusters supplemented with new genetic data (RG_R 3-7) remained similar to this observed earlier (Ozerov et al. 2017). Thus, the subdivisions to regional and local reporting groups remained the same.

Updated baseline GSI accuracy

The power of the updated baseline data was first tested using simulated mixture samples composed of a single stock only (100% simulations). Mean GSI accuracy across 186 baseline samples was 79%, which was 1 percent point (pp) higher compared to the previous baseline and varied from 4% to 100% (Table S1). Similar to the previous baseline version low GSI accuracies were observed for the populations along the southern Kola Peninsula in the White Sea, and within the Ponoï river system in Russia as well as for the populations in the Western Finnmark and Troms-Nordland regions in Norway. However, a considerable increase of the GSI accuracy was observed when samples were assigned to 26 local (consisting on average of 7 populations, mean accuracy 90%, range 7–100%) or 7 regional RGs (consisting on average of 26 populations, mean accuracy 97%, range 65–100%).

Although the overall GSI accuracies were similar to those observed earlier (Ozerov et al. 2017), a substantial increase of GSI accuracy was observed among the populations with updated genetic profiles, where the GSI accuracy increased by 6 pp for assignments to the population and RG_L and by 1 pp for assignments to RG_R . The most evident increase of accuracy was observed among the populations of the Varanger Peninsula and the R. Teno mainstem, by 2 and 3 pp, respectively.

As expected, the mean GSI accuracy across 186 baseline populations decreased when baseline performance was evaluated using hold-out mixtures of 1000 individuals randomly removed from the baseline to the mixture; a method reducing the baseline sample size by 6.7% on average. Compared to 100% simulations, the largest decreases in the GSI accuracies were observed at the population level with a mean reduction by 20 or 21 pp using the Bayesian or CML approach, respectively. However, the decreases in the mean GSI accuracies were more moderate at the RG levels: 14 and 8 pp with Bayesian, and 16 and 9 with CML method for the local and regional RGs , respectively. In general, the overall level of GSI accuracies for hold-out mixtures in the updated baseline were similar to those observed earlier (Ozerov et al. 2017).

GSI accuracy for populations of interest

Populations with the largest estimated stock sizes (categories 4 and 5, $n = 20$) were selected for more thorough evaluation of GSI success. GSI accuracy of resampled individuals from the largest stocks was highly variable across populations and was affected by the composition of the mixture sample and the level of genetic divergence (Table 7). Generally, population-level GSI accuracy was high (mean = 82%, range = 49–100%) for the regional fishery scenario, i.e. when the contribution from populations of interest to the mixture was relatively high (4–24%). In this scenario, 12 populations showed $\geq 80\%$ GSI accuracy at population level. For the non-regional fishery scenario, i.e. when the contribution from populations of interest to the mixture was low (0.5–1.5%), population-level GSI accuracy was slightly reduced (mean = 78%, range = 38–100%). However, the population-level GSI accuracy remained high ($\geq 80\%$) for 10 populations.

Given their genetic divergence, the most accurate identification of population of origin, both in regional (87–100%) and in non-regional (97–100%) mixture scenarios, was achieved for the large stocks from the Eastern Barents and White Sea basins (002-Verkhnyaya Pechora, 007-Mezen Pizhma, 008-Kuloy Soyana and 009-Severnaya Dvina Vorykva). High GSI accuracies (regional fishery scenario: 82–95%, non-regional fishery scenario: 80–91%) were also observed for 065-Kola, 144-Alta and 166-Målselv as well as for the populations of Teno (095-Teno-Iesjohka, 096-Teno-Karasjohka and 099-Teno-Inarjohka).

Table 7. GSI accuracy for 20 populations of interest based on resampled individuals with different proportions of simulated individuals in the mixtures.

Sample	ID	Reporting group		Identification of resampled individuals in mixtures with simulated individuals	
		RG _L	RG _R	Regional fishery scenario	Non-regional fishery scenario
Pechora	001UNY	1	1	87/100/100	97/100/100
Mezen	007MPI	2	1	100/100/100	99/100/100
Kuloy	008SOY	2	1	100/100/100	100/100/100
Severnaya Dvina	009VOR	2	1	100/100/100	100/100/100
Varzuga	023VRG	4	2	84/94/100	73/92/97
Strelna	027STR	4	2	57/91/100	51/79/94
Ponoi main stem	035PMS	5	2	65/84/100	58/73/86
Iokanga	047IOK	6	3	77/95/98	68/86/89
Rynda	055RYN	6	3	49/80/97	38/70/75
Kola	065KOL	7	3	81/81/98	79/79/83
Vestre Jakobselv	087VJa	10	4	75/85/86	62/67/67
Teno-Iesjohka	095IES	12	5	93/95/99	85/92/95
Teno-Karasjohka	096KARAS	12	5	86/100/100	94/98/98
Teno-Inarijohka	099Inari	13	5	90/94/99	88/88/91
Teno main stem	113TMSYK	15	5	71/93/95	75/77/82
Lakselva Porsanger	132LP	18	6	78/79/96	80/80/89
Repparfjordelva	140RF	19	6	76/77/85	72/72/79
Alta	144ALT	20	6	89/93/100	81/81/89
Reisa	154REI	21	6	87/88/97	79/79/90
Målselv	166ME	23	7	97/97/99	81/81/81

Estimated GSI accuracies at the RG levels were markedly higher. For example, for the regional fisheries scenario, reliable identification with GSI accuracy $\geq 80\%$ was achieved for 19 (mean GSI accuracy = $92 \pm 6\%$) and for all 20 populations (mean GSI accuracy = $98 \pm 3\%$) to local and regional RGs, respectively. For the non-regional fisheries scenario adequate identification with GSI accuracy $\geq 80\%$ was achieved for 10 populations to local RGs (mean GSI accuracy = $84 \pm 11\%$) and for 17 populations to regional RGs (mean GSI accuracy = $88 \pm 10\%$). In comparison with the previous baseline version, GSI accuracy increase in regional scenario was observed for 027-Strelna by 11/3/0 pp, populations of Teno (095-Teno-Iesjohka by 3/1/1 pp, 096-Teno-Karasjohka by 4/5/3 pp and 099-Teno-Inarijohka by 3/2/-1 pp), 144-Alta by 3/5/0 pp, 144-Reisa by 5/5/2 pp and 166-Målselv by 2/2/0 pp at population, RG_L and RG_R levels, respectively. Whereas for the other populations of interest the GSI accuracy was not changed or decreased. In non-regional fishery scenario, compared to the previous data (Ozerov et al. 2017), consistent GSI accuracy increase at population, RG_L and RG_R levels was observed in eight populations: 008-Kuloy Soyana, 027-Strelna, 047-Iokanga, 055-Rynda, 096-Teno-Karasjohka, 113-Teno mainstem, 132-Lakselva Porsanger and 140-Repparfjordelva. It should be noted, however, that direct comparison of the previous and current GSI simulations might be convoluted given that allele frequencies of 30 baseline populations were updated and generation of exactly the same datasets (of randomly resampled baseline samples and simulated mixtures) used for tests performed earlier is impossible.

Accuracy of stock contribution estimates for populations of interest

For populations of interest, the accuracies of estimated stock contributions to regional fishery scenario mixtures were high and in most cases varied within 1% of the actual stock proportion (176 of 200 replicates; Fig. 5). The only exception to this were stock contribution estimates for 035-Ponoi mainstem (actual: 8.6%, estimated: 9.7–10.5%).

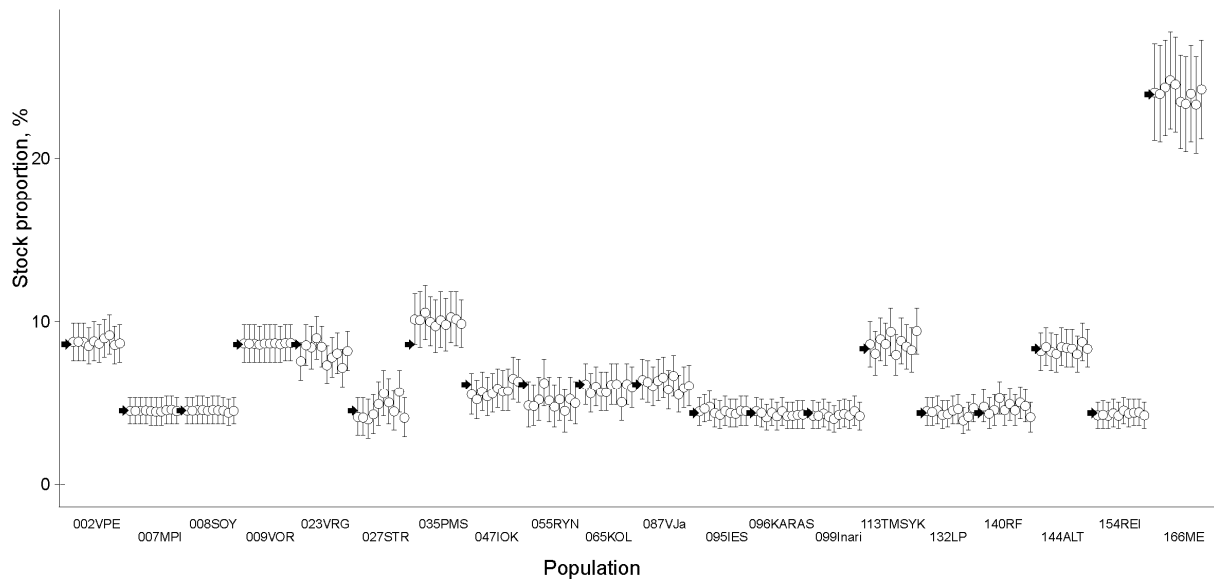


Fig. 5. Estimated stock contributions (open dots) with 95% CIs (error bars) for 20 populations of interest analyzed according to regional fishery scenario for 10 randomized datasets. Arrows indicate the actual stock proportion in the mixture. Population IDs follow those given in Table S1.

However, all actual stock proportions fall within estimated 95% CI, except for 035-Ponoi mainstem, where stock proportions were overestimated in three of 10 replicates, and for 23-Varzuga and 55-Rynda, where stock proportions were underestimated in two and three of 10 replicates, respectively (Fig. 5). In general, the accuracy of stock proportion estimates was improved for populations of interest supplemented with new samples (87-Vestre Jakobselv, 113-Teno mainstem, 132-Lakselva Porsanger, 140-Repparfjordelva, 144-Alta, 154-Reisa, 166-Målselva), compared to the previous results (Ozerov et al. 2017).

Genetic stock identification of fishery samples

In general, stock proportions in each fishery area were more similar among years within each period rather than between periods within years (Fig. 6 – Fig. 11), most likely reflecting seasonal migration patterns of different salmon populations. Catch compositions during the first period were more diverse and consisted of various populations from wide geographical area, whereas during the second period local salmon populations contributed to the fishery at larger extent.

Sør-Varanger fishery area

In Sør-Varanger fishery area we observed similar stock proportions in the fishery samples across the years, with higher contribution of the rivers of eastern coast of the Varanger fjord, i.e. from Grense Jakobselv to Vesterelva (RG_L 9), in the beginning of the fishing season (Fig. 12, 13, 15, 16). During the last weeks of the fishing season the contribution of salmon populations from the rivers of the eastern Varanger Peninsula, i.e. from Bergebyelva to Komagelva (RG_L 10), tended to increase (Fig. 12, 13, 15, 16). One exception of this pattern is the catch composition of the year 2021, where we observed exceptionally high contribution of eastern Varanger P. populations (RG_L 10; nearly 50%) to the fishery during the last weeks of the fishing season compared to the previous years (ca. 15% - 32%). However, this pattern might be, at least partly, attributed to the absence of the samples from the easternmost fishery locations of Sør-Varanger in 2021. In addition, salmon populations from the northern Kola Peninsula including Kola Bay (RG_L 6-8) substantially contributed to the Sør-Varanger fishery during the whole fishing season in 2008 – 2012 (28.6% - 39.6%), whereas the contribution of northern Kola P. populations was lower in 2020 – 2021 (12.4% - 16.3%). The largest contributors to Sør-Varanger fishery in 2008 – 2021 were salmon populations of the R. Kola (4.2% - 20.8%), R. Grense Jakobselva (9.1% - 13.3%), R. Neiden (4.9% - 30.6%) and R. Bergebyelva (5.1% - 30.9%; Table 8; Table S5).

Contribution of populations of the R. Tana system (including tributaries) to Sør-Varanger fishery varied from 7.2% to 10.2% in different years (Table 8; Table S5).

Table 8. Proportion of salmon from the rivers showing the largest contribution in Sør-Varanger fishery area.

<i>Year</i>	<i>Kola</i>	<i>Grense Jakobselva</i>	<i>Neiden</i>	<i>Bergebyelva</i>	<i>Tana</i>
2008	0.136	0.122	0.177	0.101	0.081
2009	0.112	0.000	0.306	0.051	0.102
2011	0.208	0.075	0.105	0.172	0.080
2012	0.208	0.104	0.137	0.125	0.076
2020	0.065	0.133	0.229	0.142	0.074
2021	0.042	0.113	0.049	0.309	0.072

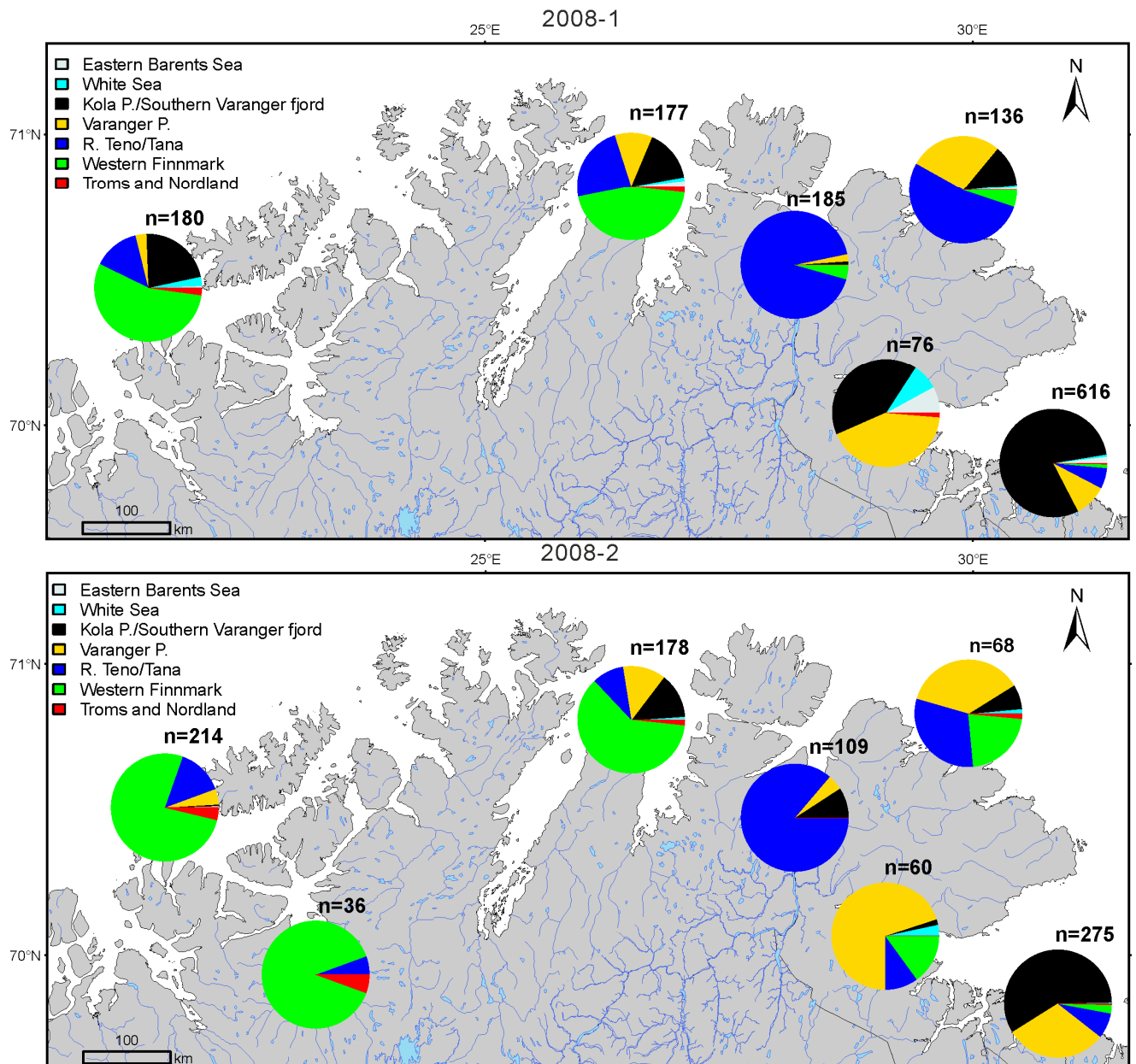


Fig. 6. Proportions of regional reporting group contributions to fishery samples in 7 fishery areas in 2008 and two periods (Table 3). Period 1 includes samples collected in June and period 2 in July – beginning of August.

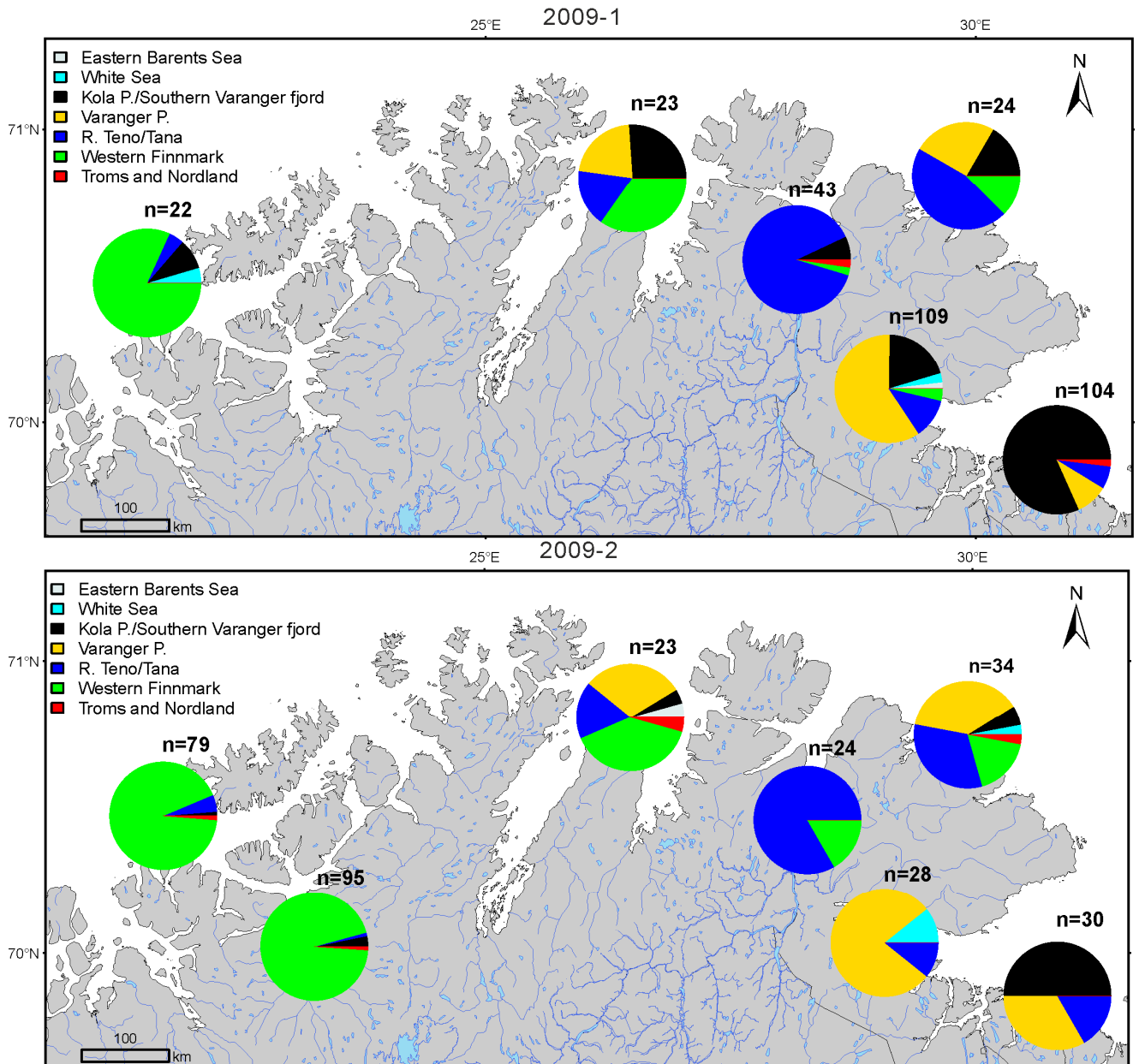


Fig. 7. Proportions of reporting group contributions to fishery samples in 7 fishery areas in 2009 and two periods (Table 3). Period 1 includes samples collected in June and period 2 in July – beginning of August.

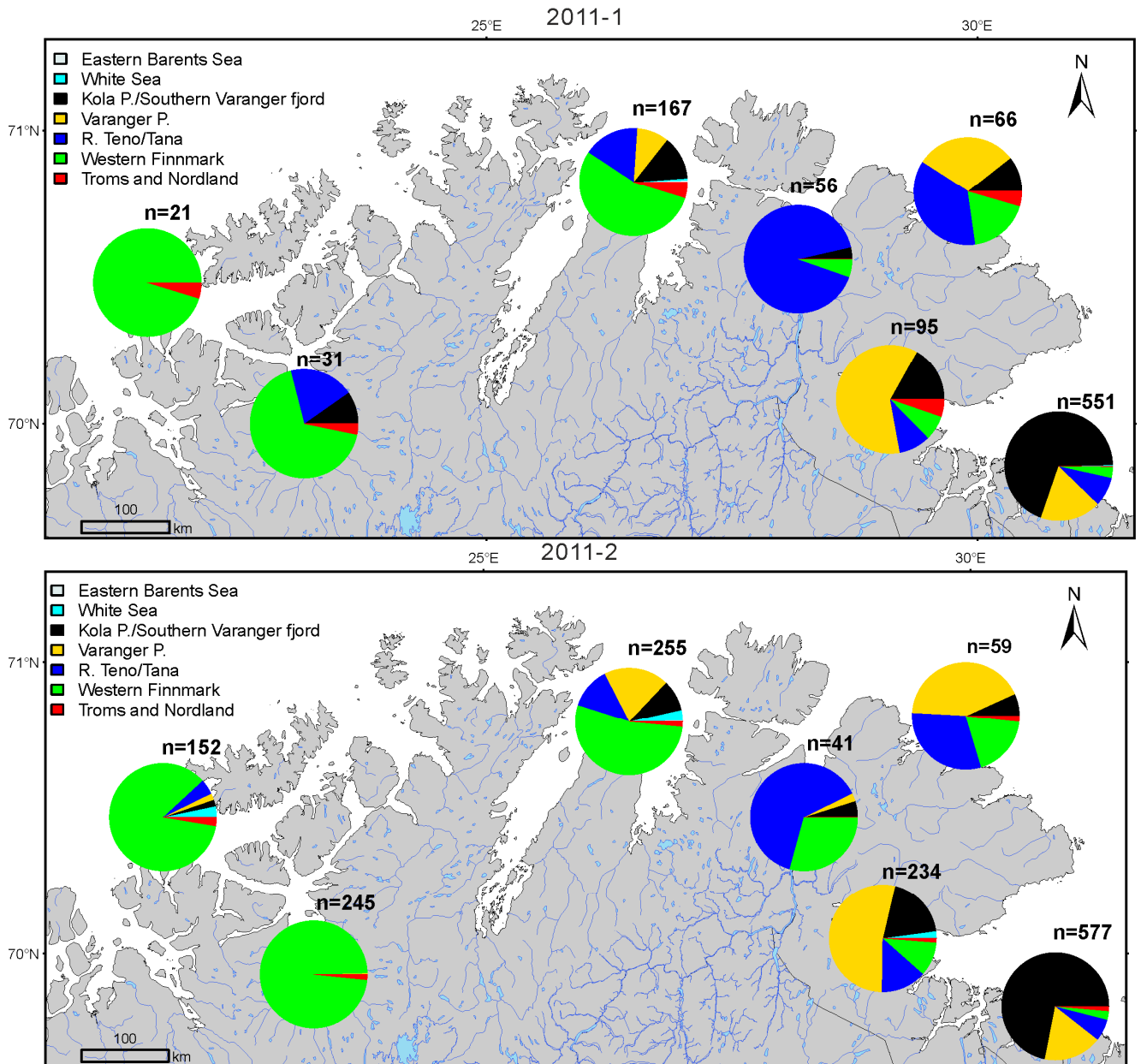


Fig. 8. Proportions of regional reporting group contributions to fishery samples in 7 fishery areas in 2011 and two periods (Table 3). Period 1 includes samples collected in June and period 2 in July – beginning of August.

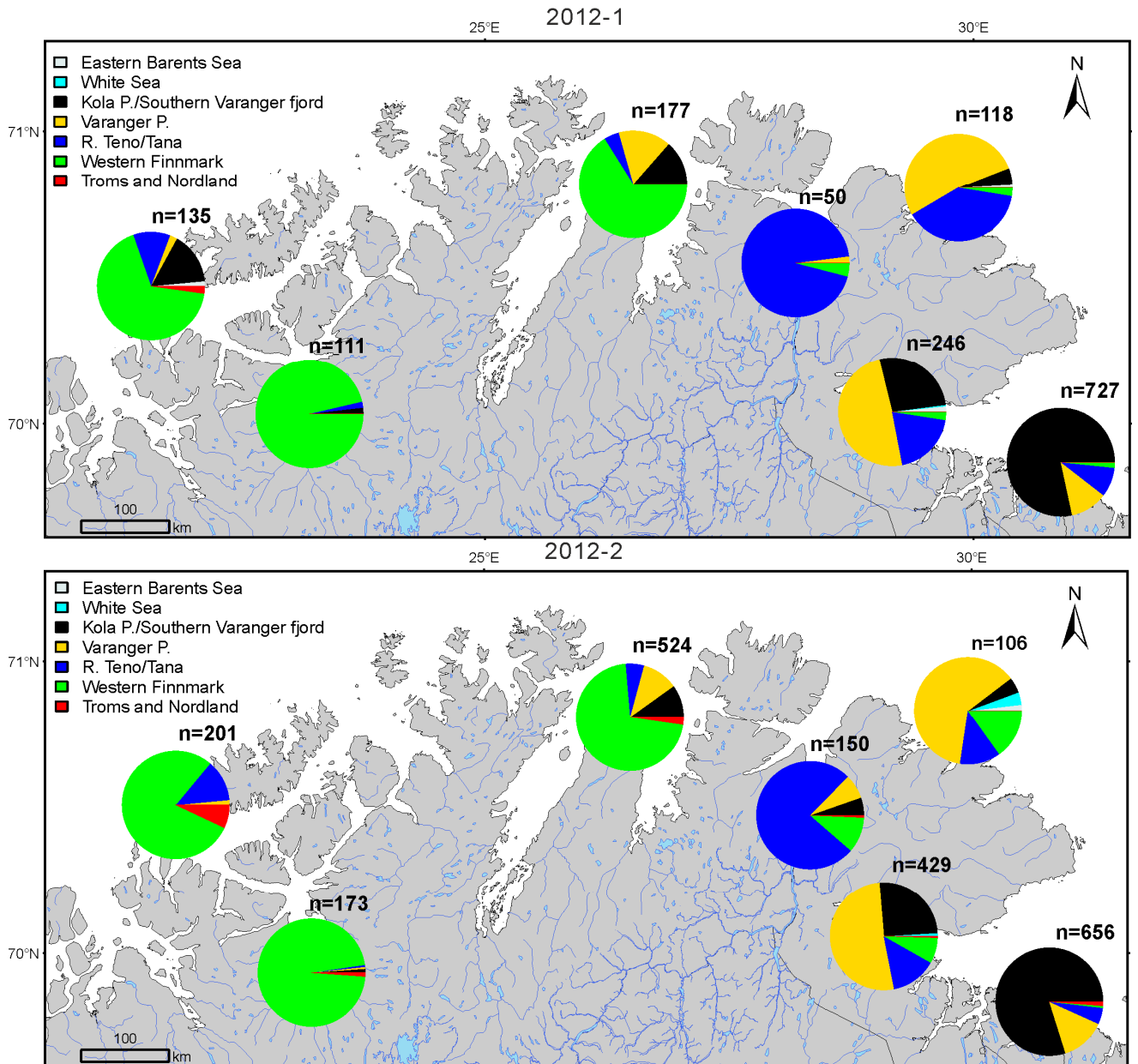


Fig. 9. Proportions of regional reporting group contributions to fishery samples in 7 fishery areas in 2012 and two periods (Table 3). Period 1 includes samples collected in June and period 2 in July – beginning of August.

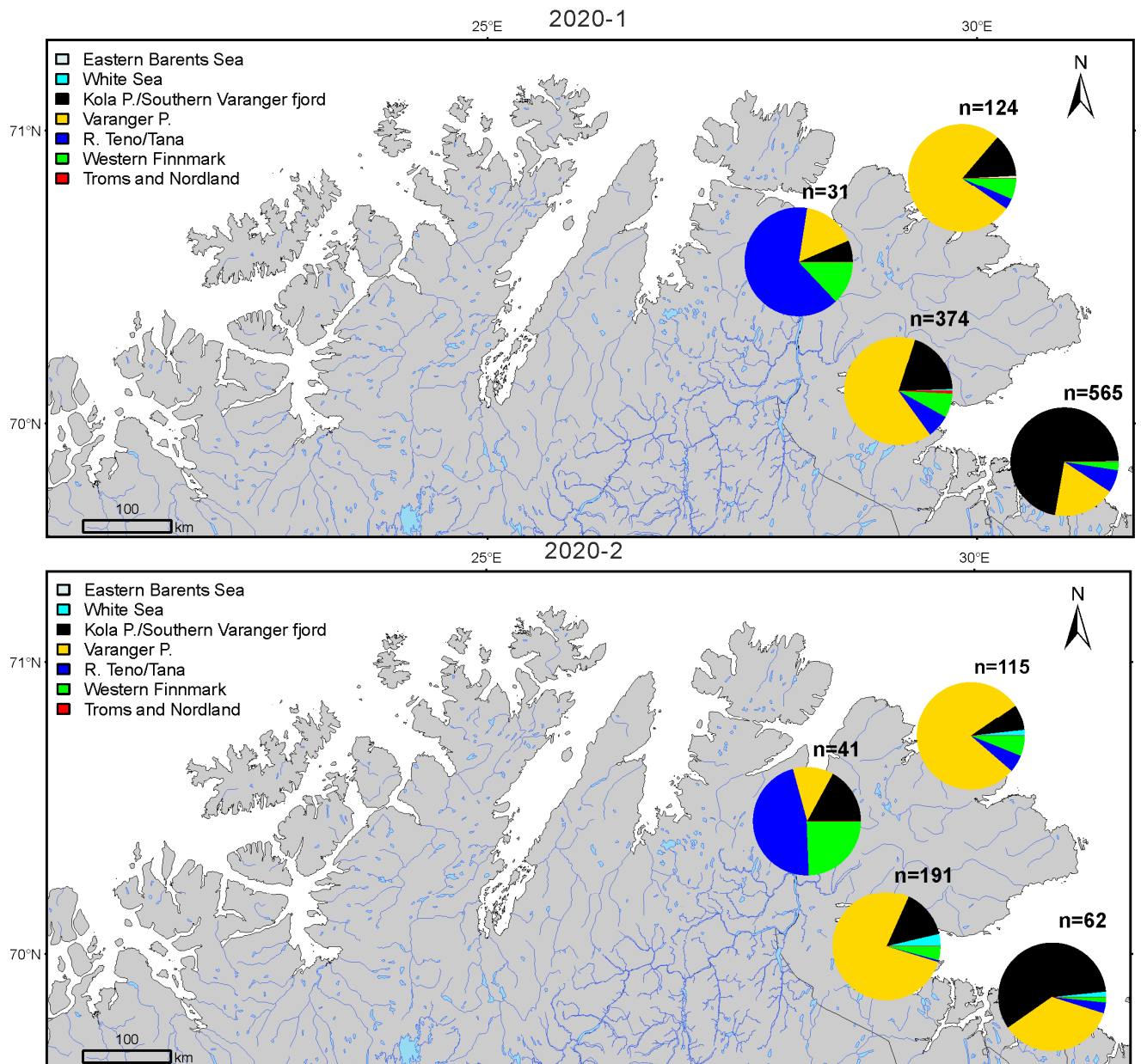


Fig. 10. Proportions of regional reporting group contributions to fishery samples in 4 fishery areas in 2020 and two periods (Table 3). Period 1 includes samples collected in June and period 2 in July – beginning of August.

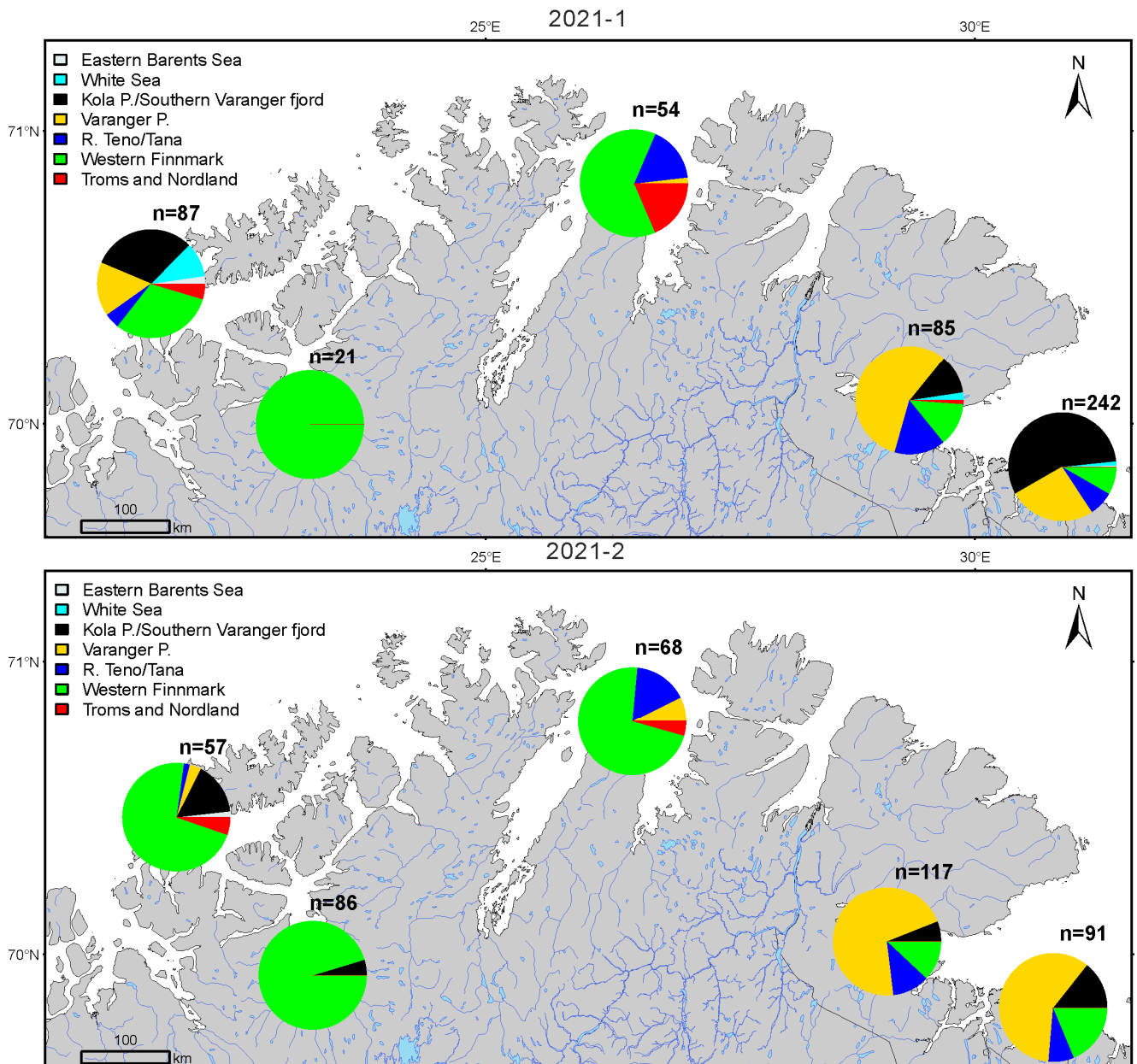


Fig. 11. Proportions of regional reporting group contributions to fishery samples in 5 fishery areas in 2021 and two periods (Table 3). Period 1 includes samples collected in June and period 2 in July – beginning of August.

Nesseby and Vadsø fishery area

The majority of fishery samples caught in Nesseby and Vadsø were composed of the fish originating from the Varanger P. rivers, mostly from the eastern part of the peninsula that includes rivers from Bergebyelva to Komagelva (RG_L 10; 43.3% - 62.0%). During the first few weeks of the fishing season various stocks contributed to the fishery, including populations from the Kola P. and R. Tana (Fig. 12, 13, 15, 16). The largest contributors to Nesseby and Vadsø fishery in 2008 – 2021 were salmon populations of the R. Bergebyelva (27.5% - 49.4%) and R. Vestre Jakobselva (6.1% - 17.8%), summing up in total from 35.7% to 61.0% in different years (Table 9). Contribution of the R. Tana system stocks to Nesseby and Vadsø fishery varied from 5.0% to 16.1% in different years (Table 9).

Table 9. Proportion of salmon from the rivers showing the largest contribution in Nesseby and Vadsø fishery area.

<i>Year</i>	<i>Bergebyelva</i>	<i>Vestre Jakobselva</i>	<i>Tana</i>
2008	0.291	0.137	0.051
2009	0.432	0.178	0.110
2011	0.280	0.144	0.140
2012	0.275	0.082	0.161
2020	0.427	0.079	0.050
2021	0.494	0.061	0.128

Vardø, Båtsfjord, Berlevåg and Gamvik fishery area

In 2008 – 2012 stock composition in Vardø, Båtsfjord, Berlevåg and Gamvik fishery area was mainly presented by relatively even contribution of salmon from northern Varanger P. rivers (RG_L 11) and populations of R. Tana system (Fig. 12, 13, 15, 16). Whereas in 2020 the proportion of Tana populations in catches was very low (Fig. 14). The proportion of the populations originating from the rivers of Varanger P. tended to increase during the second half of the fishing season (July-beginning of August; Fig. 12, 13, 15, 16). The major contributors to the fishery in Vardø, Båtsfjord, Berlevåg and Gamvik were populations from the rivers Komagelva (0.0% - 12.0%), Syltefjordelva (4.8% - 31.9%) and Kongsfjordelva (2.3% - 35.6%), which accounted from 17.5% to 68.1% to total catch in different years (Table 10). It should be noted, that the proportion of local populations in the Vardø, Båtsfjord, Berlevåg and Gamvik fishery tended to increase between 2008 and 2020. On the other hand, we observed a drastic decline of Tana populations proportion in catches of this area, which decreased nearly 10 times, from 47.4% in 2008 to 4.7% in 2020 (Fig. 14, Table 10).

Table 10. Proportion of salmon from the rivers showing the largest contribution in Vardø, Båtsfjord, Berlevåg and Gamvik fishery area.

<i>Year</i>	<i>Komagelva</i>	<i>Syltefjordelva</i>	<i>Kongsfjordelva</i>	<i>Tana</i>
2008	0.082	0.070	0.023	0.474
2009	0.071	0.048	0.167	0.381
2011	0.000	0.266	0.055	0.312
2012	0.120	0.279	0.049	0.240
2020	0.005	0.319	0.356	0.047

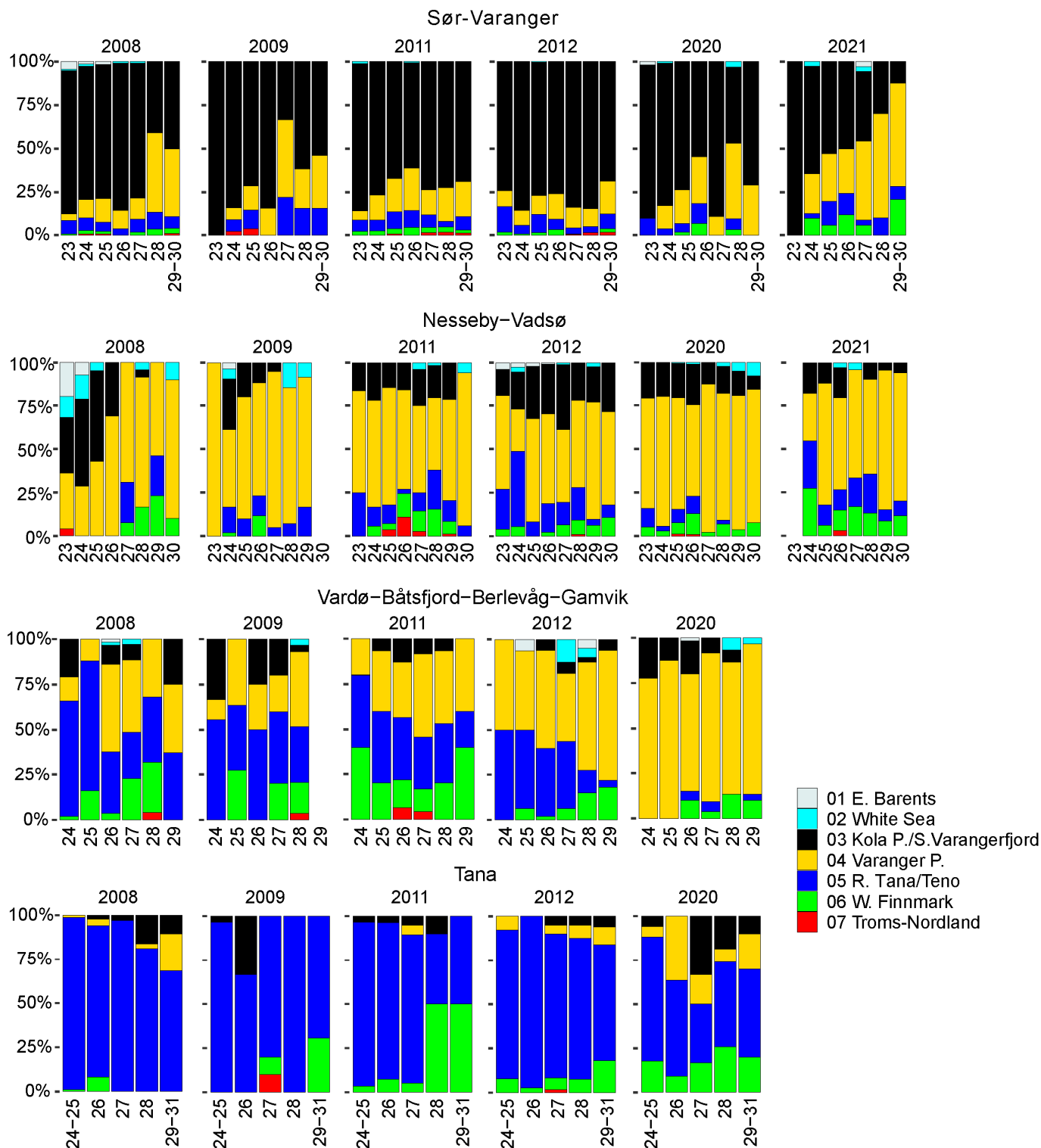


Fig. 12. Weekly contribution of regional reporting groups (RGR) to coastal fishery samples in Sør-Varanger, Nesseby-Vadsø, Vardø-Båtsfjord-Berlevåg-Gamvik and Tana fishery areas. Proportion of fish assigned to each RGR is presented at y-axis and the week number at x-axis.

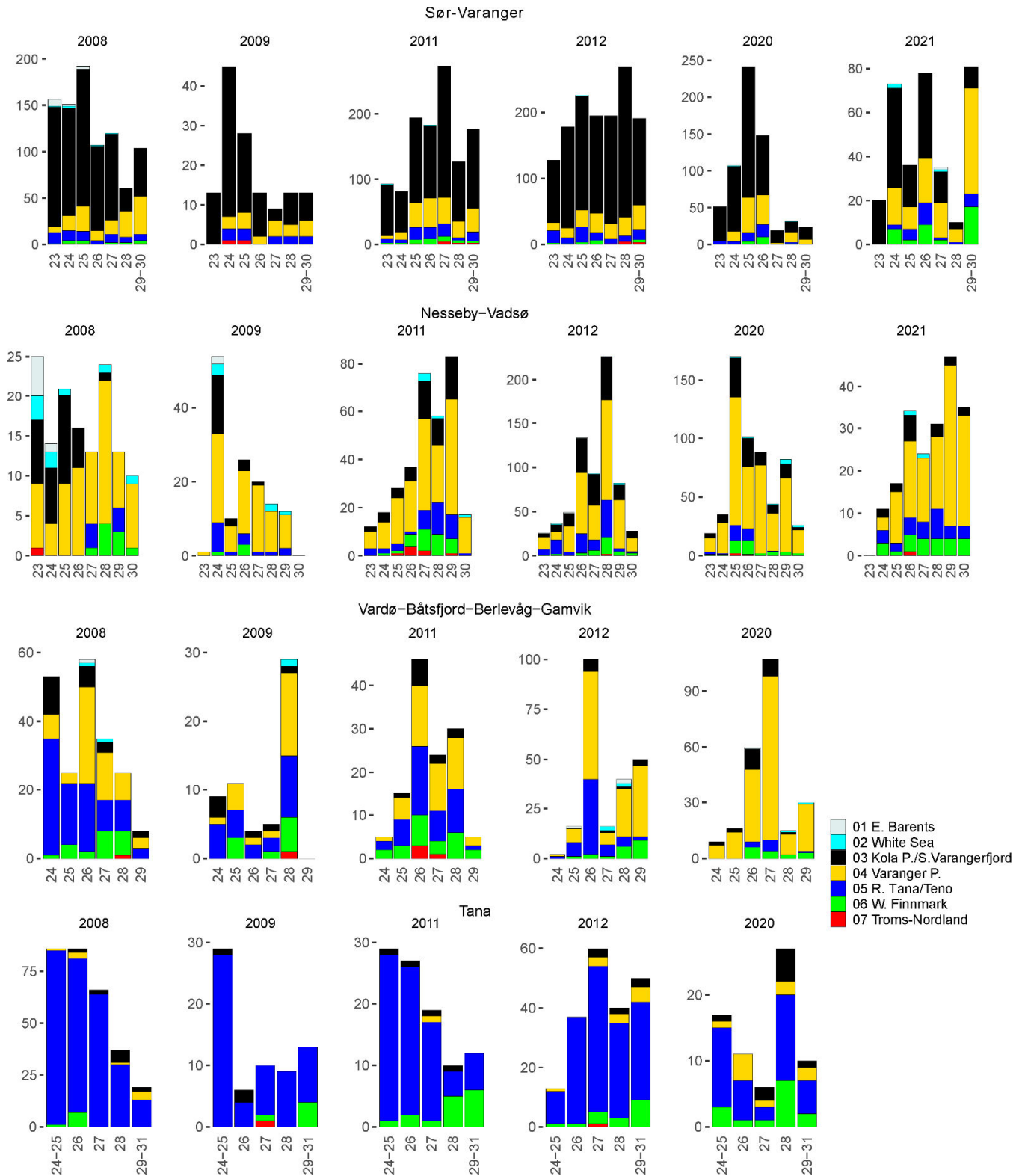


Fig. 13. Weekly contribution of regional reporting groups (RGR) to coastal fishery samples in Sør-Varanger, Nesseby-Vadsø, Vardø-Båtsfjord-Berlevåg-Gamvik and Tana fishery areas. Number of fish assigned to of each RGR is presented at y-axis and the week number at x-axis.

Tana fishery area

The majority of stocks contributed to the catches in Tana fishery area in the beginning of the season originated from the Tana river system. Later in the fishing season the growing contribution of other stocks from Western Finnmark (Lakse and Porsanger fjord), Varanger Peninsula and eastern Varanger fjord rivers was observed (Fig. 12, 13, 15, 16). However, in 2020 the overall proportion of Tana origin fish in Tanafjord catches was ca. 60%, which was much lower compared to the previous years (ca. 80 – 90%; Fig. 14). Generally, we observed a declining trend for R. Tana populations proportion in Tanafjord catches: from 89.8% in 2008 to 62.1% in 2020 (Fig. 12 – 16, Table 11), i.e. the contribution of R. Tana populations to the fishery in this area decreased by 27.7 percent points in 2020 compared to 2008.

Table 11. Proportion of salmon from the rivers showing the largest contribution in Tana fishery area.

Year	Tana	Syltefjordelva	Repparfjordelva
2008	0.898	0.000	0.004
2009	0.870	0.000	0.037
2011	0.827	0.000	0.074
2012	0.784	0.056	0.049
2020	0.621	0.121	0.121

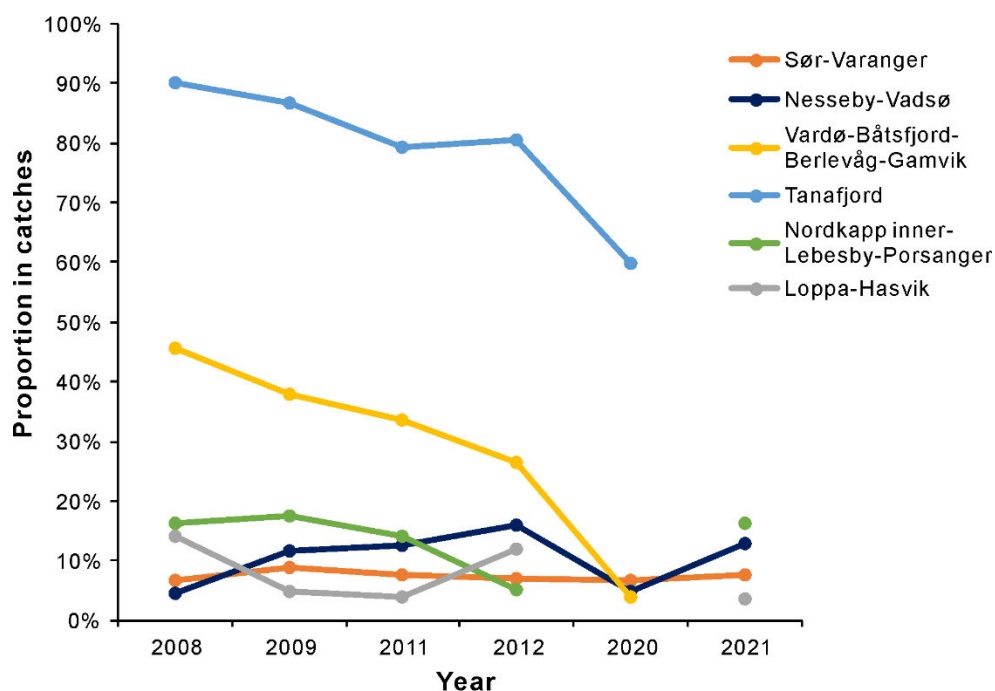


Fig. 14. Contribution of R. Tana populations to catches in seven fishery areas in 2008 – 2021.

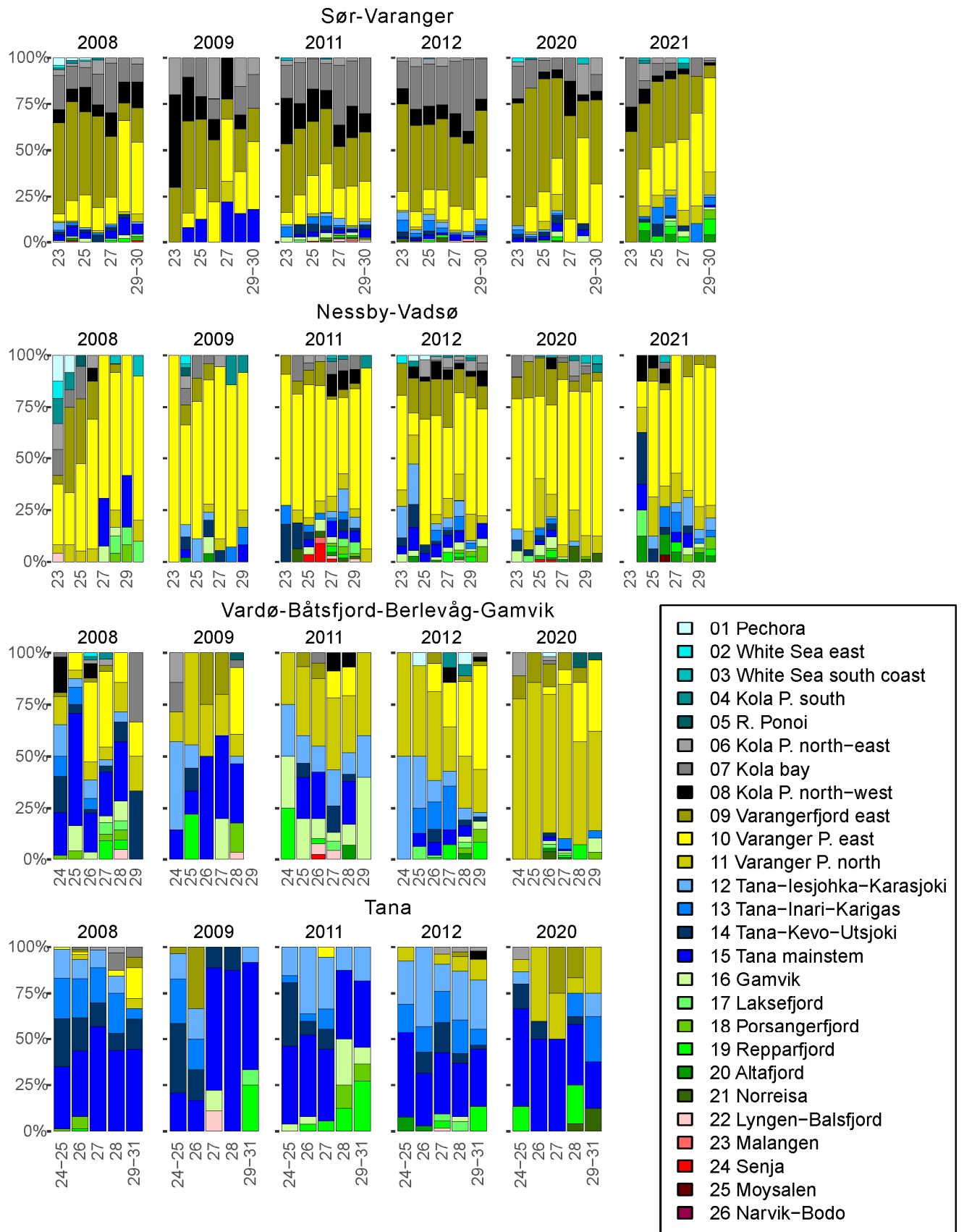


Fig. 15. Weekly contribution of local reporting groups (RGL) to coastal fishery samples in Sør-Varanger, Nessby-Vadsø, Vardø-Båtsfjord-Berlevåg-Gamvik and Tana fishery areas. Proportion of fish assigned to each RGL is presented at y-axis and the week number at x-axis.

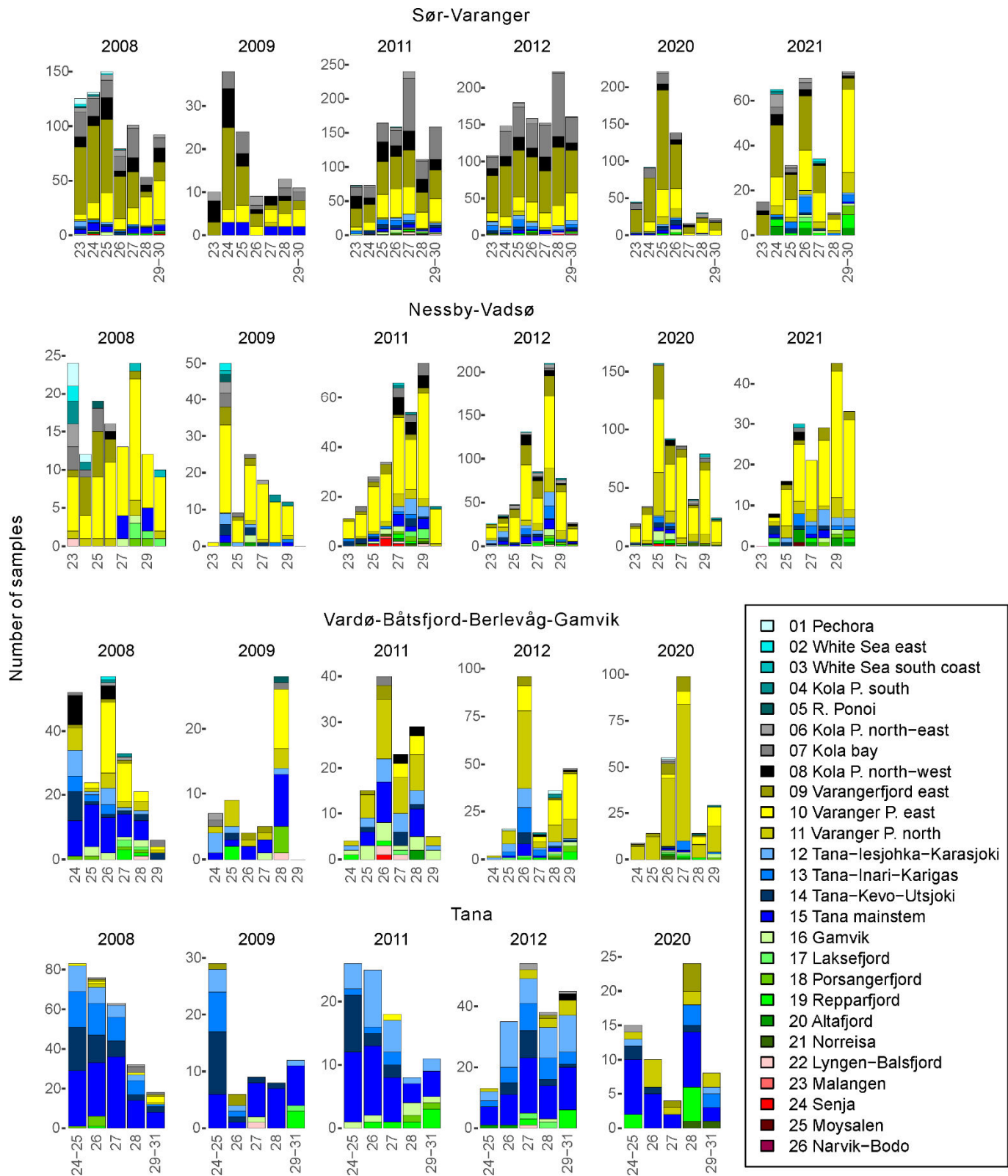


Fig. 16. Weekly contribution of local reporting groups (RGL) to coastal fishery samples in Sør-Varanger, Nessby-Vadsø, Vardø-Båtsfjord-Berlevåg-Gamvik and Tana fishery areas. Number of fish assigned to each RGL is presented at y-axis and the week number at x-axis.

Nordkapp, Lebesby and Porsanger fishery area

The samples from Nordkapp, Lebesby and Porsanger fishery were mostly composed of populations from Western Finnmark with additional contribution of various populations from Varanger P. rivers, southern Varanger fjord and R. Tana, which was substantial in some years. It should be noted, that proportion of local populations from the rivers inflowing to Lakse (RG_L 17), Porsanger (RG_L 18) and Repparfjordelva (RG_L 19) fjord increased later during the fishing season (Fig. 17 – 20). The major contributors to the Nordkapp, Lebesby and Porsanger fishery in different years were populations from the rivers Repparfjordelva (7.0% - 15.6%), Lakselva Porsanger (2.3% - 23.0%), Børselva (2.0% - 16.3%), Tana (5.4% - 18.6%), Syltefjordelva (4.1% - 15.0%) and Vestre Jakobselva (0.0% - 11.6%; Table 12).

Table 12. Proportion of salmon from the rivers showing the largest contribution to Nordkapp, Lebesby and Porsanger fishery area.

<i>Year</i>	<i>Vestre Jakobselva</i>	<i>Syltefjordelva</i>	<i>Tana</i>	<i>Børselva</i>	<i>Lakselva Porsanger</i>	<i>Repparfjordelva</i>
2008	0.000	0.068	0.162	0.020	0.111	0.105
2009	0.116	0.047	0.186	0.163	0.023	0.070
2011	0.003	0.150	0.135	0.024	0.109	0.156
2012	0.022	0.088	0.054	0.052	0.260	0.111
2021	0.000	0.041	0.175	0.093	0.144	0.103

Loppa-Hasvik fishery area

In the first half of the season the fishery samples in Loppa and Hasvik fishery area were composed of populations from the Western Finnmark with contribution of various populations from R. Tana, Kola P., Varangerfjord and Varanger P. (Fig. 17 – 20). During the second half of the fishing season local populations from the Western Finnmark were dominating in the catches (Fig. 17 – 20). R. Alta population contribution to the fishery samples in the area was major, varying from 21.6% to 47.5% in different years, followed by Repparfjordelva (5.0% - 28.6%), Tana (2.9% - 13.8%) and Lakselva Porsanger (1.7% - 12.1%; Table 13).

Table 13. Proportion of salmon from the rivers showing the largest contribution to Loppa and Hasvik fishery area.

<i>Year</i>	<i>Tana</i>	<i>Lakselva Porsanger</i>	<i>Repparfjordelva</i>	<i>Alta</i>
2008	0.138	0.120	0.156	0.216
2009	0.033	0.121	0.286	0.429
2011	0.029	0.072	0.194	0.475
2012	0.114	0.059	0.190	0.339
2021	0.042	0.017	0.051	0.297

Alta fishery area

The fishery samples in the inner Alta fjord were mostly composed of population from R. Alta (77.8% - 93.5%) being relatively stable during the fishing season and across the years (Fig. 17 – 20, Table 14). The contribution of R. Tana in Alta fjord fishery was very low and varied from 5.6% in 2008 to 0.0% in 2021.

Table 14. Proportion of salmon from the rivers showing the largest contribution to Alta fishery area.

<i>Year</i>	<i>Alta</i>	<i>Tana</i>
2008	0.778	0.056
2009	0.830	0.011
2011	0.935	0.023
2012	0.927	0.004
2021	0.933	0.000

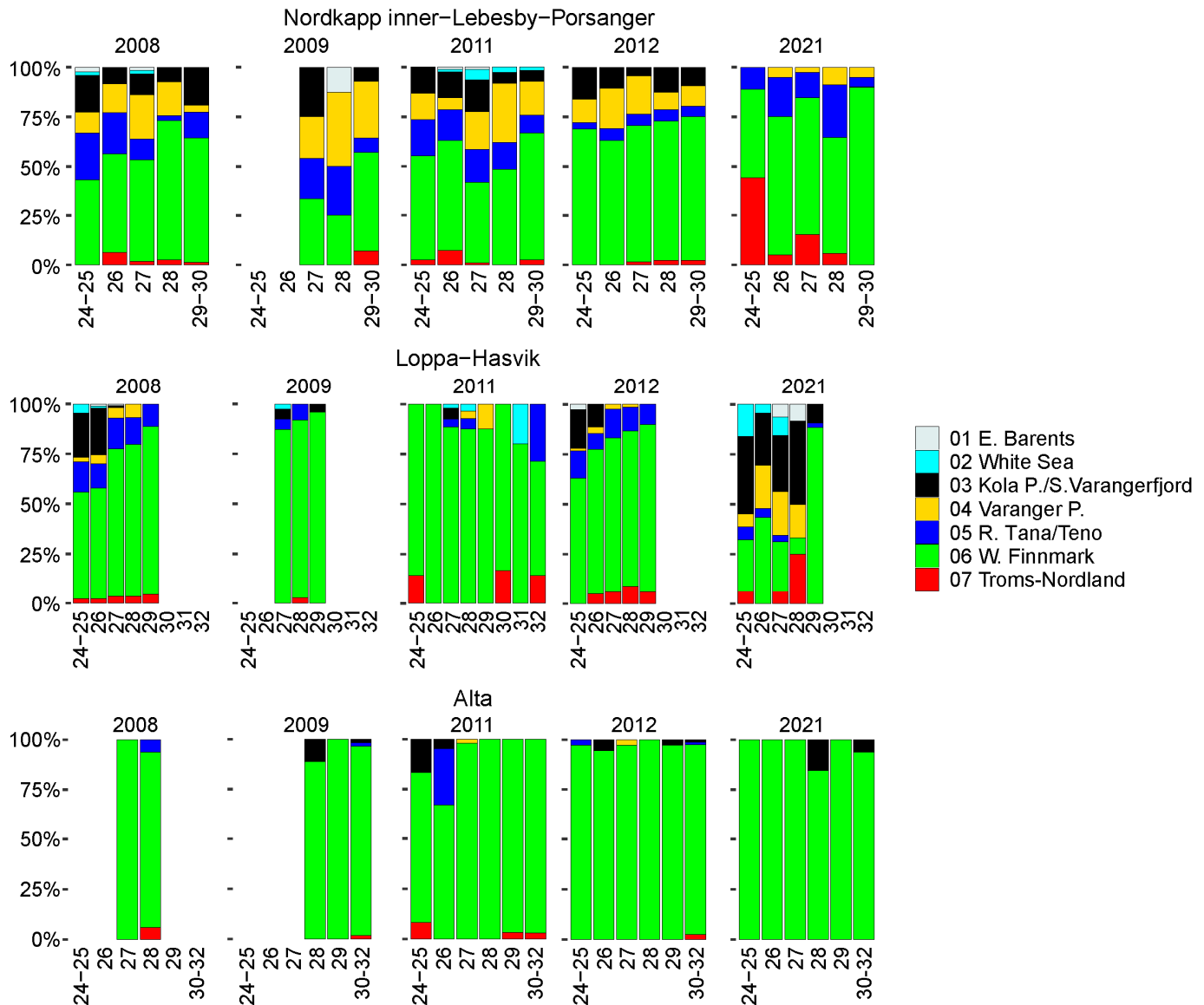


Fig. 17. Weekly contribution of regional reporting groups (RGR) to coastal fishery samples in Nordkapp inner-Lebesby-Porsanger, Loppa-Hasvik and Alta fishery areas. Proportion of fish assigned to each RGR is presented at y-axis and the week number at x-axis.

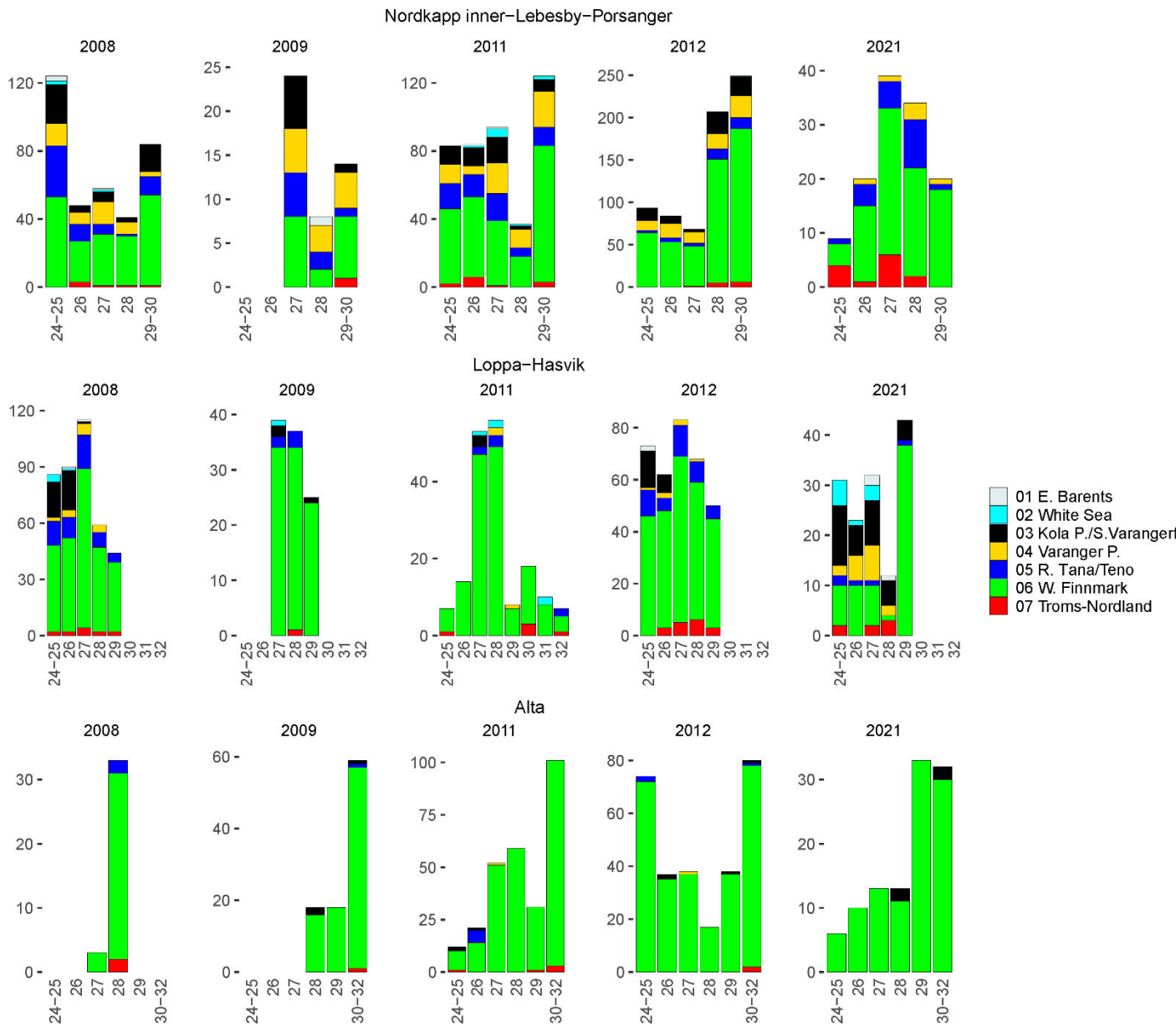


Fig. 18. Weekly contribution of regional reporting groups (RGR) to coastal fishery samples in Nordkapp inner-Lebesby-Porsanger, Loppa-Hasvik and Alta fishery areas. Number of fish assigned to each RGR is presented at y-axis and the week number at x-axis.

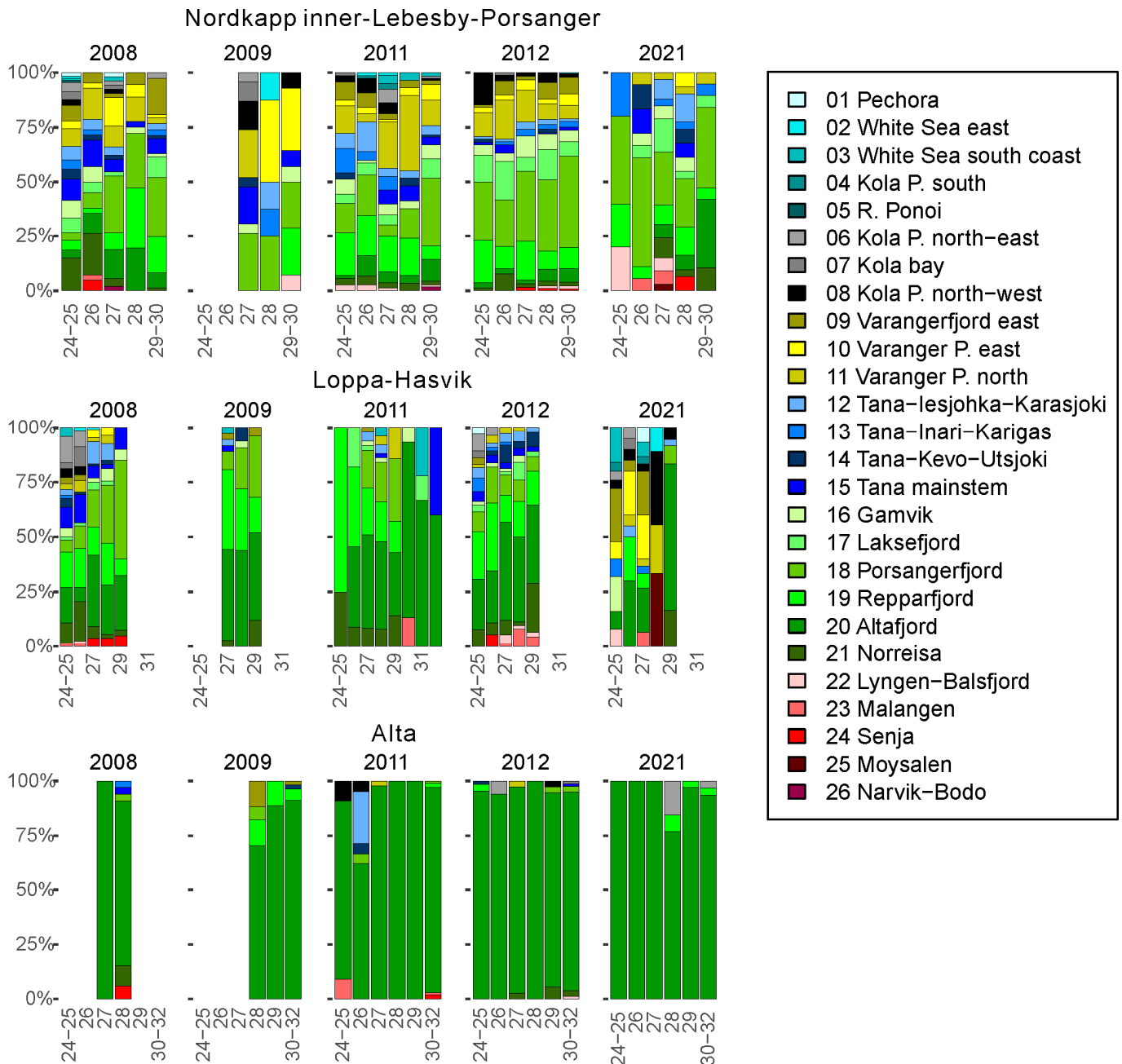


Fig. 19. Weekly contribution of local reporting groups (RGL) to coastal fishery samples in Nordkapp inner-Lebesby-Porsanger, Loppa-Hasvik and Alta fishery areas. Proportion of fish assigned to each RGL is presented at y-axis and the week number at x-axis.

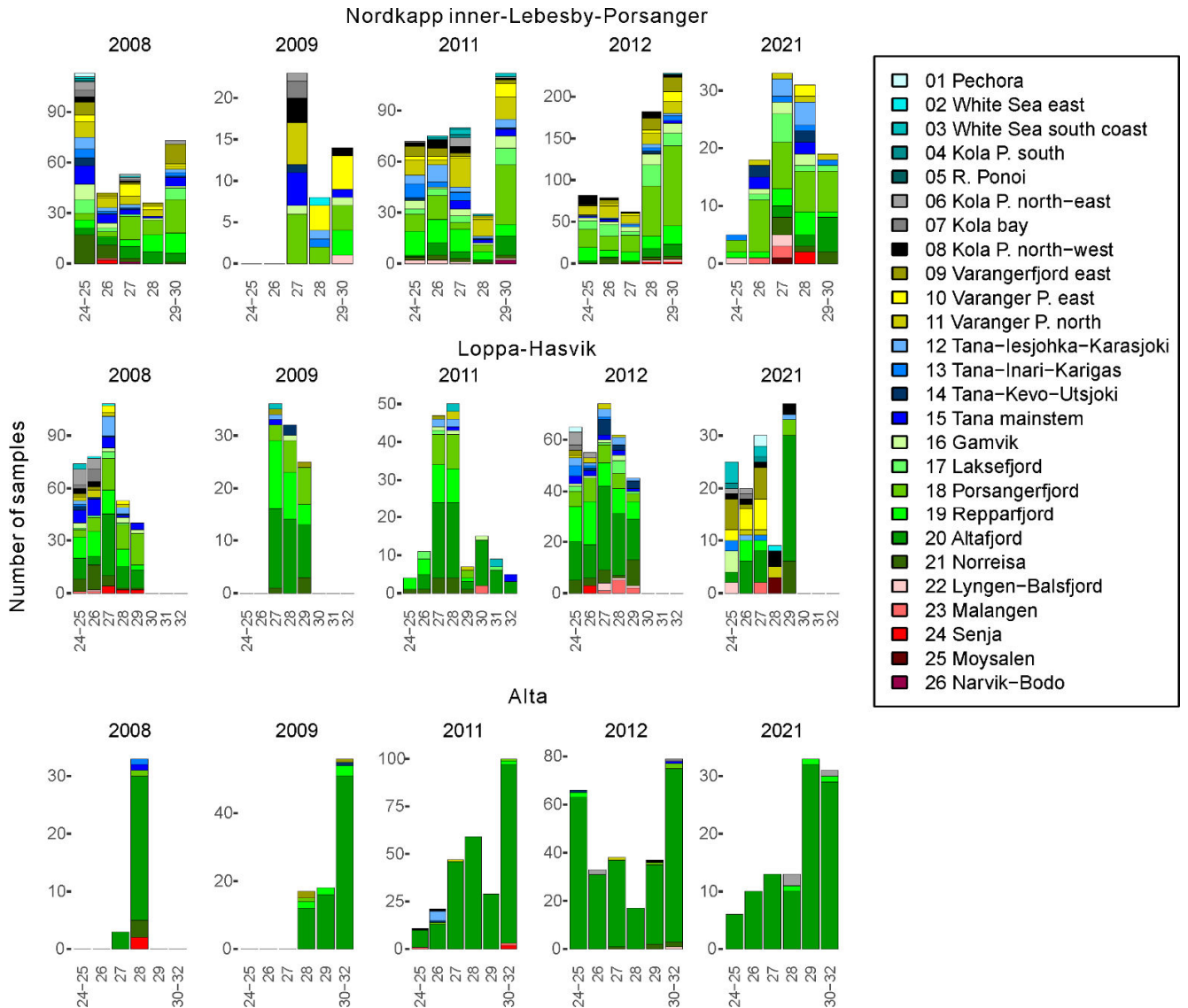


Fig. 20. Weekly contribution of local reporting groups (RGL) to coastal fishery samples in Nordkapp inner-Lebesby-Porsanger, Loppa-Hasvik and Alta fishery areas. Number of fish assigned to each RGL is presented at y-axis and the week number at x-axis.

Discussion

Worldwide decline of Atlantic salmon stock abundance during the last decades requires urgent measures directed to mitigation of the major factors affecting salmon mortality at sea, such as fishery. Given that adult salmon migrates in mixed-stocks, coastal fishery harvests a variety of populations on their way to natal rivers. Therefore, development and improvement of genetic tools for identification of the origin of salmon caught at sea to reveal the level of exploitation of certain populations is of high priority. In this report, we present the updated genetic baseline for identification of salmon caught at sea and the level of contribution of various salmon populations in northern Norwegian fishery revealed using GSI in recent and previous years (from 2008 to 2021).

Genetic structure and temporal variation of the updated baseline

We observed stable genetic structure and genetic variation among the populations in the updated baseline. The major and minor genetic clusters remained the same as in the previous baseline version (Ozerov et al. 2017). In addition, the level of genetic variation among populations remained similar to what was observed earlier. This pattern was supported by low level of genetic variation due to temporal component among the studied populations. A relative temporal stability of genetic variation in the northernmost European range of Atlantic salmon distribution has been observed earlier (Glover et al. 2012, Ozerov et al. 2013), contrasting higher levels of temporal genetic variation in southern populations of Norway (Skaala et al. 2006, Glover et al. 2012). Most likely, relative temporal stability of the northernmost populations is shaped by lower level of anthropogenic pressure in the area and, at greater extent, much lower density of salmon aquaculture in northern Norway compared to the south, although aquaculture rapidly spreads northward during the recent years (Skaala et al. 2006, Glover et al. 2012, Glover et al. 2019).

GSI power of the updated baseline

Various simulation approaches to test the power of the updated baseline showed that overall GSI accuracies were similar to those observed earlier (Ozerov et al. 2017). However, a substantial increase of GSI accuracies was observed among the populations with updated genetic profiles. Similarly, the accuracy of stock proportion estimates was largely improved for large populations supplemented with new samples, compared to the previous results (Ozerov et al. 2017). GSI success rate depends on a range of interacting factors, such as the number of baseline populations, the number and the level of polymorphism of analyzed loci, genetic divergence among populations, and baseline sample sizes (e.g., Hansen et al. 2001, Kalinowski 2004, Griffiths et al. 2010, Beacham et al. 2011). Previous studies showed that the GSI success rate was largely determined by the interaction of population genetic divergence and baseline sample size (e.g., Ozerov et al. 2017, Vähä et al. 2017). Moreover, the effect of increasing baseline sample size was higher for populations of low genetic divergence ($F_{ST} < 0.030$; Ozerov et al. 2017). Indeed, genetic divergence of the majority of populations with the updated genetic profiles (22 of 29) was low ($F_{ST} < 0.030$) as well as the average genetic divergence of these populations ($F_{ST} = 0.028$). Thus, increasing baseline sample size by 80 (30 – 134) individuals per population on average allowed improving mean GSI accuracy by up to 6 percent points at population level. In addition, variation of stock proportion estimates for large salmon populations with updated genetic profiles was low and remained within 1% of the actual stock proportion. On the other hand, we observed a slight decrease of GSI accuracy in some other populations compared to the previously published results (Ozerov et al. 2017). However, given that the genetic profiles of 30 populations were updated, and test mixtures were generated by random baseline resampling, obtaining exactly the same datasets used for the tests performed earlier is awkward. Thus, straight comparisons of the results of GSI accuracy simulation tests between this and the earlier study might be convoluted. Nevertheless, given that the levels of genetic variation and population genetic structure remained the same, which corresponds to observed relative temporal stability, we believe that the updated baseline improves GSI accuracy for populations in Varanger fjord area and for the largest rivers in the region.

GSI of fishery samples

We observed a relatively stable spatial variation in catch composition of the analyzed samples between years (2008 – 2021). In the beginning of the fishing season in June (period 1) population composition of catches was more diverse, whereas a larger proportion of local populations was contributing to the fishery during the second half of the fishing season in July – beginning of August (period 2). It should be noted, that in the majority of fishery areas annual changes of *RG* contributions to the fishery were relatively small and may largely reflect both differences in sampling sites distribution (Table 3) and variations in recruitment rate and in salmon populations' migration time.

For example, in Sør-Varanger fishery we did not observe drastic variations in the proportions of different *RG_R* and *RG_L* contributions during most of the observed years for two periods, except period 2 in 2021. However, this pattern was most likely shaped by the absence of the samples from the easternmost fishery locations in July-August of 2021. Similarly, we observed ca. 2 times decrease of salmon from northern Kola P. rivers (RR. Pechenga - Lumbovka) in 2020 (16.3%) and 2021 (12.4%) in catches, compared to 2008 – 2012 (28.6% - 39.6%), which, at least partly, may be explained by much smaller number of samples collected in the eastern part of Sør-Varanger in 2020 and 2021 compared to the previous years. On the other hand, we cannot completely exclude drastic changes or annual variations in the recruitment rates of populations from the north-western Kola P. during the recent decade. For example, an increased mortality of returning adults was observed in the Rivers Kola and Tuloma since 2015 (most likely caused by ulcerative dermal necrosis disease) and in 2020-2021 the density of juveniles in these rivers was very low compared to the previous years (S. Prusov, personal communication). Continuous catch composition monitoring of Norwegian commercial fishery in the eastern Sør-Varanger and of scientific fishery in the Russian waters of the Barents Sea (westwards Rybachy P. – Kola Bay) as well as monitoring of juvenile densities in the rivers of the Kola P. would assist to disentangle current salmon population dynamics in the area.

The fishery composition in the inner Varanger fjord (Nesseby-Vadsø fishery area) was stable over the studied years and was mostly presented by the local populations of the south-eastern part of the Varanger P. and at lesser degree of the eastern Varanger fjord. The proportion of transient salmon from other *RGs* was higher during the first weeks, but remained at low levels most of the fishing season, indicating major role of local populations in contribution to the inner fjord fishery. The proportion of northern Kola P. salmon in Nesseby-Vadsø catches was ca. 3.5 – 4.5 times lower than that observed in Sør-Varanger catches. This pattern was in line with the observations of Svenning et al. (2019) showing that the majority of R. Kola salmon (>90%) is caught close to the Varanger Fjord in eastern Finnmark due to its limited coastal movements in North-Norwegian waters. It should be noted that contribution of northern Kola P. populations to Nesseby-Vadsø fishery also decreased in 2020 (4.2%) and 2021 (2.7%) compared to the previous years (7.8% - 11.4% in 2008-2012), which is hardly explained by sampling bias in this fishery area, given good sample sizes in both Nesseby and Vadsø municipalities during all years and periods (Table 3).

The major part of catch composition in Vardø, Båtsfjord, Berlevåg and Gamvik fishery area consisted of Varanger P. and R. Tana populations (ca. 74% - 84%). However, in contrast to the Sør-Varanger and Nesseby-Vadsø fishery, the contribution of these two reporting groups was not stable over the years. For example, in 2008 and 2009 Varanger P. and R. Tana salmon contributed to the fishery nearly equally, but in 2011 and 2012 the proportion of Tana salmon decreased by 10-20 pp and in 2020 the proportion of Tana fish in the catches drastically dropped to 4%, indicating 10 times decrease since 2008. At the same time, the contribution of populations of northern and eastern Varanger P. increased in 2.5 times from 30.6% in 2008 to 77.8% in 2020 and the majority of fish caught in Vardø, Båtsfjord, Berlevåg and Gamvik fishery area in 2020 originated from the northern Varanger P. rivers (65.3%). Thus, such sharp increase of northern Varanger P. rivers contribution to the catches in this area requires re-consideration of fishery management strategy to avoid overexploitation of salmon populations of small census size.

The fishery in Tana fjord was mostly exploiting R. Tana salmon, however, the proportion of R. Tana populations in catches decreased by 27.7 percent points in 2020 compared to 2008, from 89.8% to 62.1%, which was in line with the observations of decreasing proportion of Tana fish in the catches in adjacent municipalities (Vardø, Båtsfjord, Berlevåg and Gamvik) during the studied period. The long-term observations of the number of salmon entering Tana fjord and lower part of the R. Tana corroborate our GSI results, showing similar declining trend of the number of spawners, which was the lowest in 2020 (Thorstad et al. 2021). Although the probability of catching R. Tana origin fish is the highest close to the R. Tana mouth and in areas closest to the Tana Fjord, R. Tana salmon also extensively migrates along the coast to both the east and west of the Tana fjord (Svenning et al. 2019). According to Svenning et al. (2019) ca. 40% of Tana salmon were captured >300 km west and 25% were captured >250 km east of Tana fjord. Thus, Tana stock was the most abundant and probably the most exploited by coastal salmon fishery in northern Finnmark during the recent decades, which resulted in the observed recent declines. Fishery management decided to close salmon fishery in Tana fjord and in the R. Tana, as well as in Nordkapp, Lebesby, Gamvik and Berlevåg municipalities since 2021 until further notice. This should prevent further decline of Tana populations due to fishery, however, the stock recovery may take long time, given 6-8 years' generation time of salmon populations habitating in the north.

The majority of salmon caught in Nordkapp, Lebesby and Porsanger fishery area originated from the rivers of Western Finnmark, however, the contribution of stocks from R. Tana, Varanger P., eastern Varanger fjord and Troms-Nordland was substantial, particularly during the first period of the fishing season, whereas the proportion of north-western Kola P. salmon was low. In addition, in some years salmon originating from the White Sea and eastern Barents Sea basin rivers were observed in very low numbers. This pattern was most likely driven by the fact that the majority of fishery sampling sites were located close to the mouth of the Porsanger and Lakse fjords, and only in 2012 and 2021 three and one sampling fishery sites, respectively, were located in the inner Porsanger fjord. For example, in 2009 the samples were collected only from the inner Nordkapp fishery, which was located close to the mouth of the Porsanger fjord. RG_R and RG_L composition of this fishery sample was very diverse with high proportion of salmon from the regions other than Western Finnmark. In contrast, in 2012 and 2021 we observed higher proportions of salmon originated from Western Finnmark, particularly from Porsanger fjord rivers, which was expected given the presence of samples from the inner Porsanger fjord fishery. Nevertheless, the contribution of local populations was higher during the second period of the fishing season. Thus, fishery in the mouth of the fjords targets more salmon from other regions, particularly during the first period of the fishing season. It should be also noted, that the proportion of Tana salmon in Nordkapp, Lebesby and Porsanger fishery area observed in 2021 (17.5%) was similar to that observed in 2008 (16.2%) and 2009 (18.6%), corroborating earlier observations on extensive coastal migration of the R. Tana populations (Svenning et al. 2019).

The fishery composition in Loppa and Hasvik fishery area was mostly presented by the populations of Western Finnmark, with higher diversity of RG_R and RG_L composition during the first period and larger proportion of local salmon during the second period. An exception was fishery composition of 2021, where we observed a substantial contribution of eastern Varanger fjord and Varanger P. populations during the first period of the fishing season. The overall pattern of catch composition in Loppa and Hasvik indicates that outer fjord fishery intercepts more non-local fish, particularly during the first period, similar to the fjord mouth fisheries. The contribution of the R. Alta salmon to Loppa and Hasvik fishery was the largest, being ca. 35% on average across the studied years, followed by Repparfjordelva salmon (mean = 17.5%). We expected to see large proportions of salmon from both rivers, as Loppa and Hasvik fishery operates in relatively close proximity of their mouths (ca. 80 km from the R. Alta and ca. 100 km from Repparfjordelva). The higher proportion of the R. Alta salmon contribution to the catches is due to its larger census population size (SSC 5) compared to the R. Repparfjordelva (SSC 4, Table S1).

The most stable catch composition was observed in the inner Alta fjord fishery area, where salmon from the R. Alta was dominating (77.8% - 93.5%) in the catches during the whole fishing season and

across the years. This was in line with the stable status of R. Alta salmon stock during the recent decade (<https://lakseregisteret.fylkesmannen.no/visElv.aspx?id=212.Z>) and the observations that the majority of Alta salmon is caught along a 60 km section of the coast proximate to the Alta River (Svenning et al. 2019).

Conclusions

GSI of the adult salmon caught at sea showed that stock proportions during the fishing season in each fishery area were more similar among years within each period (June and July-August) rather than between periods within years most likely reflecting migration patterns of different salmon populations. Catch compositions during the first period (June) were more diverse and consisted of various populations from wide geographical area, whereas during the second period (July-August) local salmon populations contributed to the fishery at larger extent.

We observed relatively stable stock composition across the years within each period in Sør-Varanger, Nesseby and Vadsø, Nordkapp, Lebesby and Porsanger, Loppa and Hasvik, and Alta fishery areas. As exception, the composition of Sør-Varanger catches in 2021 was presented by larger proportion of eastern Varanger P. populations during the last weeks of the fishing season compared to the previous years, which was most likely due to the absence of samples from the easternmost fishery locations of the Sør-Varanger in 2021. In addition, salmon populations from the northern Kola Peninsula including Kola Bay substantially contributed to the Sør-Varanger fishery during the whole fishing season in 2008 – 2012, whereas the contribution of northern Kola P. populations was lower in 2020 – 2021. The most stable catch composition was observed in inner Alta fjord fishery area, where salmon from the R. Alta was dominating (77.8% - 93.5%) during the fishing season and across the years.

We observed a drastic decline of Tana salmon contribution to the fishery in Tana fjord and in adjacent municipalities (Vardø, Båtsfjord, Berlevåg and Gamvik). Thus, in 2008 – 2012 stock composition in Vardø, Båtsfjord, Berlevåg and Gamvik fishery area was mainly consisting of relatively even contribution of salmon from northern Varanger P. rivers and populations of R. Tana system. However, in 2020 the proportion of Tana populations in catches was very low, decreased nearly 10 times, from 47.4% in 2008 to 4.7% in 2020. Similarly, in 2020 the overall proportion of Tana origin fish in Tana fjord catches was ca. 60%, which was much lower compared to the previous years (ca. 80-90%). Generally, we observed a declining trend of R. Tana populations proportion in Tana fjord catches: from 89.8% in 2008 to 62.1% in 2020, which was in line with observations of decreasing number of ascending adults in Tana fjord and R. Tana system during the recent years.

We recommend continuous catch composition monitoring of Norwegian commercial fishery in northern Norway and of scientific fishery in the Russian waters of the Barents Sea as well as monitoring of juvenile densities in the rivers of northern Norway and the Kola P. to get further insights into the status and dynamics of Atlantic salmon populations in the area.

Acknowledgements

We thank professional fishers and salmon fishing organizations in Finnmark for help with providing unique biological material from the salmon catches. We thank Eevaliisa Kivilahti for assistance during fieldwork sampling and Abdullah Sami Madhun, Synne Winterthun and Johanne Øyro at IMR and Tutku Aykanat (University of Helsinki) for laboratory assistance. We acknowledge CSC-IT Center for Science, Finland for providing generous computational resources. This study was funded by the European Union, Kolarctic ENPI CBC project KO4178 “Conserving our Atlantic salmon as a sustainable resource for people in the North; fisheries and conservation in the context of growing threats and a changing environment” (CoASal). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the report. This publication has been

produced with the assistance of the European Union, but the contents can in no way be taken to reflect the views of the European Union.

Supplementary data

Supplementary file 1. Results of genetic stock identification (GSI) of Atlantic salmon caught in commercial fishery in 2008-2021 in Northern Norway.

Table S1. Studied rivers, population ID (ID), genetic diversity indices of the studied Atlantic salmon samples assessed at 31 microsatellite loci, GSI accuracy based on 100% simulation and random resampling of 1000 individuals to the mixture. SSC - stock size category; RG_L - local reporting group; RG_R - regional reporting group; n - number of samples used for analysis; H_O – observed heterozygosity; H_E – unbiased expected heterozygosity; f – within-population inbreeding coefficient; pop F_{ST} - population-specific F_{ST} , estimated as an average pairwise F_{ST} to the 20 closest neighbouring populations; N_e - effective population size and its 95% confidence interval; GSI accuracy of simulated datasets at population (POP), local (RG_L) and regional (RG_R) reporting group levels.

Table S2. The results of the Hardy-Weinberg equilibrium test across studied loci and samples. *P*-values remained significant at 1% level after correction for false discovery rate (FDR) are highlighted in bold italic.

Table S3. Estimation of null allele frequencies (Brookfield 1; Brookfield, 1996) across studied loci and samples. The frequency of a null allele significantly higher than zero are highlighted in bold.

Table S4. Pairwise genetic distances among Atlantic salmon populations based on microsatellite data as measured with F_{ST} (below diagonal) and corresponding *P*-values (above diagonal). *P*-values > 0.01 are indicated in bold.

Table S5. River contributions to coastal fishery samples in 2008 - 2021 in different areas. Proportion of each stock contributing to the fishery was calculated based on the number of fish assigned to a population with probability of ≥ 0.7 .

References

- Ackerman, M. W., C. Habicht, and L. W. Seeb. 2011. Single-Nucleotide Polymorphisms (SNPs) under Diversifying Selection Provide Increased Accuracy and Precision in Mixed-Stock Analyses of Sockeye Salmon from the Copper River, Alaska. *Transactions of the American Fisheries Society* **140**:865-881.
- Adamack, A. T., and B. Gruber. 2014. PopGenReport: simplifying basic population genetic analyses in R. *Methods in Ecology and Evolution* **5**:384-387.
- Anderson, E. C., R. S. Waples, and S. T. Kalinowski. 2008. An improved method for predicting the accuracy of genetic stock identification. *Canadian Journal of Fisheries and Aquatic Sciences* **65**:1475-1486.
- Beacham, T. D., B. McIntosh, and C. G. Wallace. 2011. A comparison of polymorphism of genetic markers and population sample sizes required for mixed-stock analysis of sockeye salmon (*Oncorhynchus nerka*) in British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences* **68**:550-562.
- Benjamini, Y., and Y. Hochberg. 1995. Controlling the False Discovery Rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**:289-300.
- Bradbury, I. R., L. C. Hamilton, G. Chaput, M. J. Robertson, H. Goraguer, A. Walsh, V. Morris, D. Reddin, J. B. Dempson, T. F. Sheehan, T. King, and L. Bernatchez. 2016a. Genetic mixed stock analysis of an interceptory Atlantic salmon fishery in the Northwest Atlantic. *Fisheries Research* **174**:234-244.
- Bradbury, I. R., L. C. Hamilton, S. Rafferty, D. Meerburg, R. Poole, J. B. Dempson, M. J. Robertson, D. G. Reddin, V. Bourret, M. Dionne, G. Chaput, T. F. Sheehan, T. L. King, J. R. Candy, and L. Bernatchez. 2015. Genetic evidence of local exploitation of Atlantic salmon in a coastal subsistence fishery in the Northwest Atlantic. *Canadian Journal of Fisheries and Aquatic Sciences* **72**:83-95.
- Bradbury, I. R., L. C. Hamilton, T. F. Sheehan, G. Chaput, M. J. Robertson, J. B. Dempson, D. Reddin, V. Morris, T. King, and L. Bernatchez. 2016b. Genetic mixed-stock analysis disentangles spatial and temporal variation in composition of the West Greenland Atlantic Salmon fishery. *ICES Journal of Marine Science: Journal du Conseil*.
- Carlsson, J. 2008. Effects of microsatellite null alleles on assignment testing. *Journal of Heredity* **99**:616-623.
- Crozier, W. W., P. J. Schon, G. Chaput, E. C. E. Potter, N. O. Maoileidigh, and J. C. MacLean. 2004. Managing Atlantic salmon (*Salmo salar* L.) in the mixed stock environment: challenges and considerations. *Ices Journal of Marine Science* **61**:1344-1358.
- Do, C., R. S. Waples, D. Peel, G. M. Macbeth, B. J. Tillett, and J. R. Ovenden. 2014. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Mol Ecol Resour* **14**:209-214.
- Earl, D. A., and B. M. vonHoldt. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**:359-361.
- Elphinstone, M. S., G. N. Hinten, M. J. Anderson, and C. J. Nock. 2003. An inexpensive and high-throughput procedure to extract and purify total genomic DNA for population studies. *Molecular Ecology Notes* **3**:317-320.
- ESRI. 2015. ArcGIS Desktop: Release 10.5.1. Environmental Systems Research Institute, Redlands, CA.
- Excoffier, L., and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**:564-567.

- Falush, D., M. Stephens, and J. K. Pritchard. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**:1567-1587.
- Gilbey, J., V. Wennevik, I. R. Bradbury, P. Fiske, L. P. Hansen, J. A. Jacobsen, and T. Potter. 2017. Genetic stock identification of Atlantic salmon caught in the Faroese fishery. *Fisheries Research* **187**:110-119.
- Glover, K. A., M. Quintela, V. Wennevik, F. Besnier, A. G. Sorvik, and O. Skaala. 2012. Three decades of farmed escapees in the wild: a spatio-temporal analysis of Atlantic salmon population genetic structure throughout Norway. *PLoS One* **7**:e43129.
- Glover, K. A., K. Urdal, T. Næsje, H. Skoglund, B. Florø-Larsen, H. Otterå, P. Fiske, M. Heino, T. Aronsen, H. Sægrov, O. Diserud, B. T. Barlaup, K. Hindar, G. Bakke, I. Solberg, H. Lo, M. F. Solberg, S. Karlsson, Ø. Skaala, A. Lamberg, Ø. Kanstad-Hanssen, R. Muladal, O. T. Skilbrei, and V. Wennevik. 2019. Domesticated escapees on the run: the second-generation monitoring programme reports the numbers and proportions of farmed Atlantic salmon in > 200 Norwegian rivers annually. *ICES Journal of Marine Science* **76**:1151-1161.
- Griffiths, A. M., G. Machado-Schiaffino, E. Dillane, J. Coughlan, J. L. Horreo, A. E. Bowkett, P. Minting, S. Toms, W. Roche, P. Gargan, P. McGinnity, T. Cross, D. Bright, E. Garcia-Vazquez, and J. R. Stevens. 2010. Genetic stock identification of Atlantic salmon (*Salmo salar*) populations in the southern part of the European range. *BMC Genetics* **11**:27.
- Hansen, M. M., E. Kenchington, and E. E. Nielsen. 2001. Assigning individual fish to populations using microsatellite DNA markers. *Fish and Fisheries* **2**:93-112.
- Hansen, M. M., E. E. Nielsen, and K. L. D. Mensberg. 1997. The problem of sampling families rather than populations: relatedness among individuals in samples of juvenile brown trout *Salmo trutta* L. *Molecular Ecology* **6**:469-474.
- Hedrick, P. W. 2005. A Standardized Genetic Differentiation Measure. *Evolution* **59**:1633-1638.
- Hill, W. G. 1981. Estimation of effective population size from data on linkage disequilibrium. *Genet Res (Camb)* **38**:209-216.
- Jakobsson, M., and N. A. Rosenberg. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**:1801-1806.
- Jones, O. R., and J. Wang. 2010. COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources* **10**:551-555.
- Jost, L. O. U. 2008. GST and its relatives do not measure differentiation. *Molecular Ecology* **17**:4015-4026.
- Kalinowski, S. T. 2004. Genetic polymorphism and mixed-stock fisheries analysis. *Canadian Journal of Fisheries and Aquatic Sciences* **61**:1075-1082.
- Kalinowski, S. T., K. R. Manlove, and M. L. Taper. 2008. ONCOR: a computer program for genetic stock identification, v.2. . Available from <http://www.montana.edu/kalinowski/Software/ONCOR.htm>.
- Keenan, K., P. McGinnity, T. F. Cross, W. W. Crozier, and P. A. Prodöhl. 2013. diveRcity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution* **4**:782-788.
- Lankia, T., R. Venesjärvi, and E. Pouta. 2022. Importance-performance analysis of the fishing tourism service structure: Recreational anglers' preferences on the remote salmon river of Teno in Finland. *Fisheries Research* **254**:106425.
- Meirmans, P. G., and P. H. Van Tienderen. 2004. genotype and genodive: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* **4**:792-794.
- Neaves, P. I., C. G. Wallace, J. R. Candy, and T. D. Beacham. 2005. CBayes: computer program for mixed stock analysis of allelic data, version v5.01. Available: http://www.pac.dfo-mpo.gc.ca/sci/mgl/Cbayes_e.htm.
- O'Sullivan, R. J., M. Ozerov, G. H. Bolstad, J. Gilbey, J. A. Jacobsen, J. Erkinaro, A. H. Rikardsen, K. Hindar, and T. Aykanat. 2022. Genetic stock identification reveals greater use of an

- oceanic feeding ground around the Faroe Islands by multi-sea winter Atlantic salmon, with variation in use across reporting groups. *ICES Journal of Marine Science* **79**:2442–2452.
- Ozerov, M., E. Niemelä, V. Wennevik, A. S. Madhun, A. Vasemägi, S. Prusov, and T. Kalske. 2020. Updated guide to collect samples representative of salmon populations, and for viral and parasite-induced diseases monitoring, 4 pp.
- Ozerov, M., J.-P. Vähä, V. Wennevik, E. Niemelä, M.-A. Svenning, S. Prusov, R. Diaz Fernandez, L. Unneland, A. Vasemägi, M. Falkegård, T. Kalske, B. Christiansen, and W. S. G. Handling editor. 2017. Comprehensive microsatellite baseline for genetic stock identification of Atlantic salmon (*Salmo salar* L.) in northernmost Europe. *ICES Journal of Marine Science* **74**:2159-2169.
- Ozerov, M. Y., A. E. Veselov, J. Lumme, and C. R. Primmer. 2013. Temporal variation of genetic composition in Atlantic salmon populations from the Western White Sea Basin: influence of anthropogenic factors? *BMC Genetics* **14**:88.
- Pella, J., and M. Masuda. 2001. Bayesian methods for analysis of stock mixtures from genetic characters. *Fishery Bulletin* **99**:151-167.
- Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**:945-959.
- R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rodriguez-Ramilo, S. T., and J. Wang. 2012. The effect of close relatives on unsupervised Bayesian clustering algorithms in population genetic structure analysis. *Mol Ecol Resour* **12**:873-884.
- Rosenberg, N. A., J. K. Pritchard, and J. L. Weber. 2002. Genetic structure of human populations. *Science (New York N.Y.)* **298**.
- Skaala, O., V. Wennevik, and K. Glover. 2006. Evidence of temporal genetic change in wild Atlantic salmon, *Salmo salar* L., populations affected by farm escapees. *ICES Journal of Marine Science* **63**:1224-1233.
- Svenning, M.-A., M. Falkegård, E. Niemelä, J.-P. Vähä, V. Wennevik, M. Ozerov, S. Prusov, J. B. Dempson, M. Power, and P. Fauchald. 2019. Coastal migration patterns of the four largest Barents Sea Atlantic salmon stocks inferred using genetic stock identification methods. *ICES Journal of Marine Science* **76**:1379-1389.
- Thorstad, E. B., T. Forseth, and P. Fiske. 2021. Status for norske laksebestander i 2021. 16.
- Vähä, J.-P., J. Erkinaro, M. Falkegård, P. Orell, and E. Niemelä. 2017. Genetic stock identification of Atlantic salmon and its evaluation in a large population complex. *Canadian Journal of Fisheries and Aquatic Sciences* **74**:327-338.
- Vähä, J.-P., V. Wennevik, M. Ozerov, R. Diaz Fernandez, L. Unneland, K. Haapanen, E. Niemelä, M. A. Svenning, M. Falkegård, S. Prusov, I. Lyzhov, K. Rysakova, T. Kalske, B. Christiansen, and G. Ustyuzhinsky. 2014. Genetic structure of Atlantic salmon in the Barents region and genetic stock identification of coastal fishery catches from Norway and Russia. Kolarctic ENPI CBC—Kolarctic salmon project (KO197) Genetic report, 95 pp.
- Vähä, J. P., J. Erkinaro, E. Niemelä, and C. R. Primmer. 2007. Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. *Molecular Ecology* **16**:2638-2654.
- Vähä, J. P., J. Erkinaro, E. Niemelä, C. R. Primmer, I. Saloniemi, M. Johansen, M. Svenning, and S. Brors. 2011. Temporally stable population-specific differences in run timing of one-sea-winter Atlantic salmon returning to a large river system. *Evolutionary Applications* **4**:39-53.
- Waples, S. R. 2006. A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics* **7**:167-184.
- Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* **38**:1358-1370.
- Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2-book.org/>.

