

Blue mussels *Mytilus edulis* and *M. trossulus* in sympatry: assessment of ecological niche divergence using species distribution modeling

V. Khaitov^{1,2,*}, P. Safonov³, A. Zaichikova⁴, M. Katolikova¹, M. Ivanov¹, P. Strelkov^{1,5}

¹St. Petersburg State University, University embankment 7–9, 199034 St. Petersburg, Russia

²Kandalaksha State Nature Reserve, Lineynaya 35, 184042 Kandalaksha, Russia

³Institute of Cytology RAS, Tikhoretsky 4, 194064 St. Petersburg, Russia

⁴Moscow State University, Leninskie gory 1c12, 119234 Moscow, Russia

⁵Laboratory of Monitoring and Conservation of Natural Arctic Ecosystems, Murmansk Arctic University, Kapitana Egorova 15, 183038 Murmansk, Russia

ABSTRACT: Species distribution models (SDMs) describing the relationship between species occurrence and environmental parameters can be used to assess the ecological niche of a species. Usually applied to morphologically distinct species, SDMs are also a promising tool for describing niche partitioning in coexisting cryptic species. An example of the latter in the marine realm are blue mussels *Mytilus edulis* L. (*ME*) and *M. trossulus* Gould (*MT*). Despite considerable research effort, little is known about how they share space and resources in sympatry anywhere except the Baltic Sea. Salinity, substrate, surf and proximity to harbors have been suggested as candidate factors of segregation, but no general consensus has been reached. Here, we assessed partial effects of these predictors on divergence of *ME* and *MT* in the White Sea littoral applying SDMs to 570 mussel samples with known taxonomic structure. We found that each of the predictors influenced spatial segregation. The most expected habitat of *ME* was a bottom substrate in a wind-exposed location with an average White Sea salinity (24 psu) away from ports and large rivers, while for *MT*, it was an algal substrate in a wind-protected area with a lower salinity close to ports and large rivers. We also found that the species segregation by substrate was density-dependent: the degree of segregation positively depended on *ME* abundance, which indicates that *ME* outcompetes *MT* on bottom substrates. We suggest that the same predictors can drive the segregation of these 2 species outside the White Sea.

KEY WORDS: *Mytilus* · Cryptic species · Species distribution models · Ecological niche divergence

1. INTRODUCTION


Species distribution models (SDMs) are a numerical tool describing the relationship between species occurrence and environmental parameters. They can be used to predict distribution patterns of species in space and time and to assess their ecological niche (Elith & Leathwick 2009). Joint application of SDMs to coexisting species, i.e. a community, makes it possible to describe the partitioning of their ecological niches (Ovaskainen & Abrego 2020).

In other words, SDMs may describe the axes in ecological space along which coexisting species are segregated. SDMs can be built using various approaches, from regular multiple regressions to advanced machine learning (Elith et al. 2006, Caradima et al. 2019, Poggiato et al. 2021).

SDMs are usually applied to morphologically distinct species (e.g. Reiss et al. 2011, Lindegren et al. 2022), which are distinguished in routine biodiversity assessment studies. However, there is increasing evidence about coexistence of cryptic species (Bick-

*Corresponding author: polydora@rambler.ru

ford et al. 2007, Geller et al. 2010, Struck et al. 2018) and infraspecific taxa (Dufresnes et al. 2023). Sympatry usually arises as a result of cryptic invasions (Morais & Reichard 2018). Coexisting taxa are unlikely to have identical ecological phenotypes, and an ecological niche partitioning between them can be expected (Sáez & Lozano 2005). The question how such taxa share space and resources in sympatry can be answered using SDMs (Peterson et al. 2019). This approach has already been successfully used in marine ecology (e.g. Dennis & Hellberg 2010, Lowen et al. 2019, Hu et al. 2021). The results indicate that cryptic taxa indeed have distinct ecological phenotypes. Therefore, ecological niche assessment of coexisting cryptic species, particularly those of economic, conservation or ecosystem importance, is a promising research direction.

Blue mussels (*Mytilus edulis* species complex) are the longest known and best studied cryptic species in the marine realm (Knowlton 1993, Gosling 2021). Six species that make up this complex hybridize in sympatry and are er to distinguish genetically than morphologically (Wenne et al. 2020, Gardner et al. 2021). Blue mussels are powerful ecosystem engineers and important species for aquaculture (Buschbaum et al. 2009, Gosling 2021).

The dominant species of blue mussels in the North Atlantic are *M. edulis* L. (*ME*) and *M. trossulus* Gould (*MT*). Their distribution on the oceanic scale is mostly regulated by temperature and its correlates such as sea-ice extent and primary production (Hayhurst & Rawson 2009, Wenne et al. 2020). Both of these species occur in the Arctic, but *ME* is distributed further south than *MT* in temperate seas (Wenne et al. 2020). There are multiple zones of sympatry between *ME* and *MT* (hereafter, contact zones), from Scotland and the Gulf of Maine in the south to Greenland and Spitsbergen in the north (Wenne et al. 2020). *ME* and *MT* evolved in allopatry since the Pliocene in the Atlantic and the Pacific Ocean, respectively, and their contact zones are thought to have formed as a result of repeated *MT* invasions from the Pacific Ocean to the Atlantic as well as from one part of the Atlantic into another (Väinölä & Strelkov 2011, Wenne et al. 2020 and references therein).

In contact zones, *ME*, *MT* and their hybrids are often found in the same settlements (Väinölä & Strelkov 2011, Wenne et al. 2020), which are referred to as mixed settlements. Scientists generally agree that sympatric *ME* and *MT* are ecologically distinct (Riginos & Cunningham 2005, Katolikova et al. 2016, Michalek et al. 2021) and have a different economic value in aquaculture (Penney et al. 2002, Beaumont et

al. 2008), but the data on the factors of their ecological segregation are fragmentary and contradictory.

The greatest progress in comparative ecological studies of *ME* and *MT* in sympatry has been made in the contact zones in the Baltic Sea, in the waters of the Kola Peninsula (White and Barents Seas) and in the West Atlantic (mainly Gulf of Maine and Nova Scotia). In the Baltic Sea, the brackish areas of its inner part are inhabited by *MT*, while the saltier areas closer to the North Sea are inhabited by *ME*. In the middle there is the contact zone, where mixed settlements can be dominated by hybrids, with *MT* gene frequency gradually increasing towards the inner Baltic (Väinölä & Strelkov 2011, Zbawicka et al. 2014, Stuckas et al. 2017). As a result, the species distribution is strongly correlated with salinity, and the role of other factors is negligible (Kijewski et al. 2019).

In the contact zones of the Kola Peninsula and the West Atlantic, there are few hybrids in mixed settlements, and the spatial distribution of *ME* and *MT* is patchy on different scales, from dozens of kilometers to dozens of centimeters. The relationship between the distribution and salinity is not obvious in these contact zones, and there seems to be no simple 'single-factor' driver of species distribution (Riginos & Cunningham 2005, Katolikova et al. 2016, Wenne et al. 2020, Marchenko et al. 2023). Depth, fouling substrate, anthropogenic pollution levels and surf effects have been considered, apart from salinity, as possible factors affecting the segregation of *ME* and *MT* (Bates & Innes 1995, Comesaña et al. 1999, Hellou & Law 2003, Tam & Scrosati 2014, Marchenko et al. 2023), but no consensus has been reached.

In particular, in the White and Barents Seas, the frequency of *MT* is greater in port areas, possibly because this species was introduced into the region with ship traffic in historic times (Väinölä & Strelkov 2011, Katolikova et al. 2016). The only segregation factor explicitly tested in the White Sea is the substrate of littoral mussels (Katolikova et al. 2016). Adult *MT* are more common on fucoid algae while adult *ME* mostly live directly on the bottom substrates such as mud, sand, stones and gravel. However, segregation across substrates cannot fully explain the local-scale patchiness (Katolikova et al. 2016). In the Barents Sea, no correlation with substrate has been found. However, these species have different depth preferences there: *ME* appears to be a more sublittoral species and *MT* a more littoral one (Marchenko et al. 2023).

It should be noted that some of the candidate factors of species segregation may be collinear and confound the analysis. Ports are often located in sheltered areas close to river mouths, and the effects of

shipping, surf and salinity are difficult to distinguish. The effects of depth and substrate may obscure each other since fucoids, common in the littoral, are rare in the sublittoral, where they are replaced by kelps (Druehl & Green 1982).

Comparative ecological studies of *ME* and *MT*, which began as early as the 1980s (see Riginos & Cunningham 2005 for review), have been hampered by 2 circumstances. Firstly, it was impossible to examine large amounts of material because species identification required labor-intensive genotyping methods (Khaitov et al. 2021). Secondly, until recently there were no reliable statistical methods for modeling the distribution of sympatric taxa in the space of multiple factors, and the SDM approach could not be implemented. To our knowledge, it has been applied to *ME* and *MT* only twice, in the above-mentioned studies by Kijewski et al. (2019) and Wenne et al. (2020). In both studies, machine learning techniques were used to model the macro-geographic distribution of species (technically, of allele frequencies at taxonomically informative genes) in the space of multiple climatic and oceanographic characteristics available from public databases. These methods have never been applied to modeling species distributions within contact zones.

In our previous studies, we found a simple semi-diagnostic conchological trait for *ME* and *MT* (Katolikova et al. 2016). Using this trait, one could make reliable interpretations of the taxonomic structure of mixed settlements on the basis of morphotype frequencies in samples, i.e. without genotyping. This procedure was referred to as the ‘morphotype test’ (Khaitov et al. 2021).

The aim of this study was to estimate the divergence of ecological niches between *ME* and *MT* in the White Sea littoral along environmental gradients such as substrate, salinity, surf level and distance from ports. All of these factors have been suggested as potentially influencing segregation of these 2 species in sympatry. Another candidate factor, depth (Marchenko et al. 2023), was not examined in our study but was controlled by sampling at the same littoral level. To achieve our aim, we examined the variability of the environmental predictors and the taxonomic structure of mussel settlements using extensive material (95 study sites, 570 mussel samples, 55 529 mussels) and assessed the partial influence of the predictors on the distribution of the proportion of *MT* using SDMs. Since all predictors were included in 1 model, collinearity could be controlled. Ideally, a model trained on reliable data should be transferable and work well on independent data. Therefore, we evaluated the predictive power of our model using testing data sets from the White Sea

and the Barents Sea. In addition, to reveal possible competition between the 2 species, we checked whether the pattern of their segregation by substrate was density-dependent.

2. MATERIALS AND METHODS

2.1. Study area

The study area was the Kandalaksha Bay, where all previous *ME* and *MT* studies in the White Sea have been conducted (Katolikova et al. 2016, Khaitov et al. 2018, 2023). The Bay is 185 km long and funnel-shaped, with a highly indented coastline and numerous islands and skerries (Fig. 1). Climate is continental subarctic with 4–5 mo of ice cover and an average monthly sea surface temperature in August of 13.8°C. Mean tidal range is about 2 m. Summer surface salinity is 24 psu in most of the Bay (average salinity for the White Sea) and lower in the estuarine areas (Berger & Naumov 2000). Two canals of a hydropower plant and 24 rivers with a catchment area of 141–12 830 km² (median 240 km²; see Table S1 in Supplement 1 at www.int-res.com/articles/suppl/m773pXXX_suppl1.pdf; note that all supplementary tables are in Supplement 1) flow into the Bay, with the largest river, the Niva, entering the Bay at its very top. Due to the indented shoreline and numerous rivers, local surf and salinity gradients are pronounced (Filatov et al. 2007).

Six ports operating oceanic vessels were functioning in the area in the 20th century (Fig. 1). Two of them, both at the top of the Bay, are still in operation. The other 4 have been abandoned (Sailing directions of the White Sea 1932, Krasavtsev 2011) but are occasionally visited by small ships (our observations).

In 2002–2013, both *ME* and *MT* were almost ubiquitous in the Bay, but their ratio in settlements varied greatly, with *ME* being generally dominant (Katolikova et al. 2016). Mussels in the Bay are particularly abundant in the littoral fucoid belt (mainly *Fucus vesiculosus* L. and *Ascophyllum nodosum* L.), which is continuous 0.5–1.0 m above mean spring tide depth (Berger et al. 2001).

2.2. Modeling data set

2.2.1. Mussel sampling and processing

Mussels were sampled at 95 sites within the littoral fucoid belt in summer months of 2011–2018 (Fig. 1). Data for 17 of these sites were included in the study



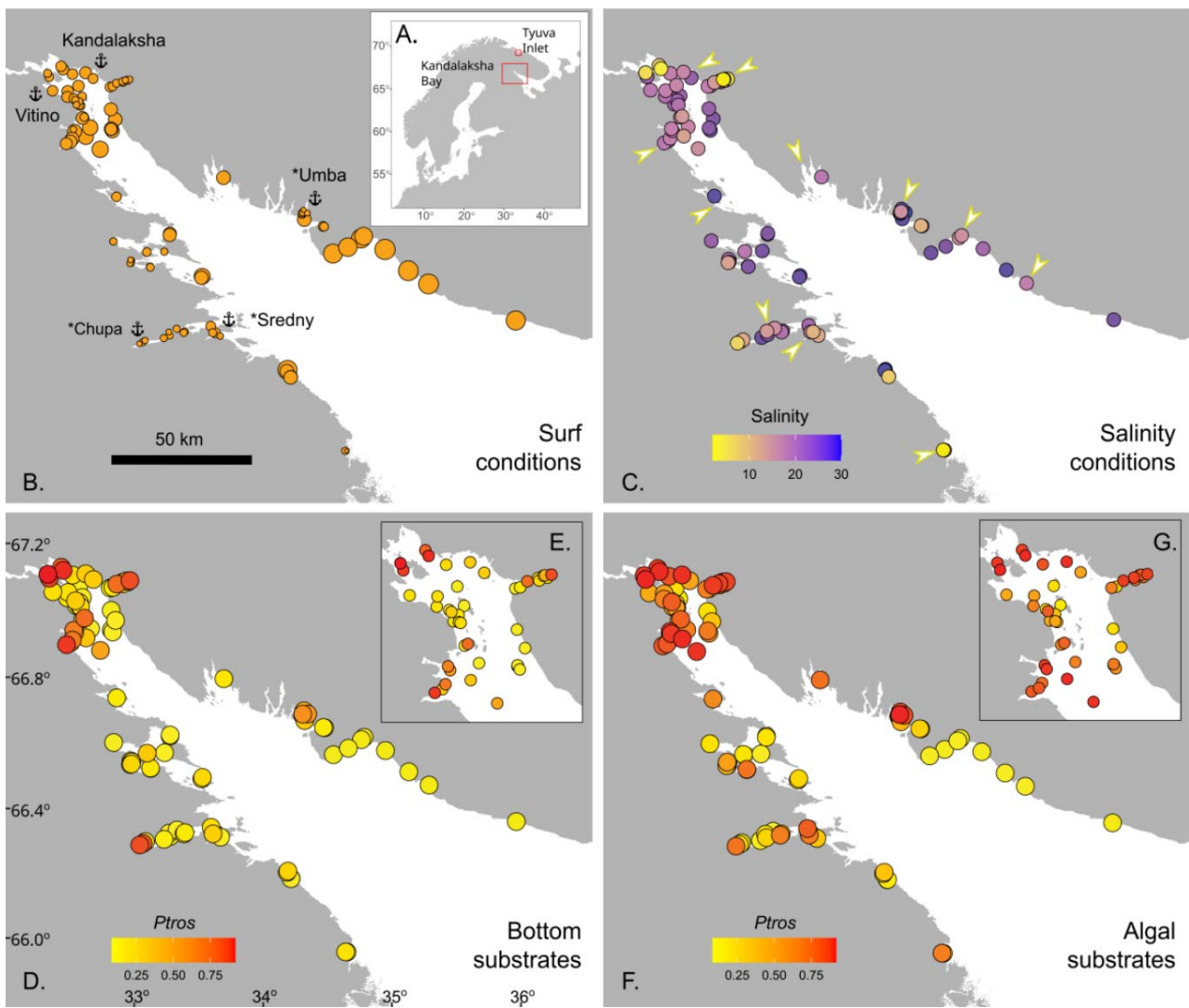


Fig. 1. Taxonomic structure of mussel settlements and their habitat characteristics. (A) Northern Europe. Red boxes mark the positions of Kandalaksha Bay and Tyuva Inlet. (B–G) Kandalaksha Bay. (B) Surf conditions. Point size is proportional to fetch. Anchors with names mark ports. Asterisks identify whether the port is currently abandoned. (C) Salinity conditions. Point filling is proportional to salinity. Arrowheads mark mouths of large rivers. (D–G) Proportion of *Mytilus trossulus* in bottom ($P_{tros_{Bottom}}$, D–E) and algal ($P_{tros_{Algae}}$, F–G) samples. Point filling is proportional to P_{tros} . Insets E and G show the Bay's top at higher resolution

by Katolikova et al. (2016); the other data are new (Table S2). The sites were chosen to describe the littoral populations of the Bay in as much detail as possible and to account for the heterogeneity of their habitat by substrate type, surf level and distance from rivers and ports. All samples were taken within the fucoid belt. At each site, 3 samples from fucoid thalli (hereafter, algal samples) and 3 samples from bottom substrates (bottom samples) were collected a few meters from each other using 0.25 m² and 0.025 m² frames, respectively. A greater size of the algal frame was associated with the large size of the fucoids and the need to account for their complex geometry. The frames were placed not randomly but approximately

at the same depth and in such a way as to capture the dense mussel aggregations.

We used mussels with a shell length larger than 10 mm for a reliable identification of the shell morphotypes (Khaitov et al. 2021). In the bottom samples, all mussels from a frame were used. In the algal samples, the procedure was different. One bundle of algae, containing at least a few dozen mussels, was chosen and weighed together with the attached mussels. The remaining algae from a frame were also weighed. Mussels from the bundle were counted and used for further analysis. The ratio between the counted number of mussels and the bundle weight was applied to the total algal weight to reconstruct

the total number of mussels in the sample (Table S3). The information on the total number of mussels in algal samples was lacking for 12 of the sites, and they were excluded from the analyses which required data on mussel abundance (Model 2 and Model 3 in Sections 2.4.1 and 2.4.2).

Mussel shell morphotypes were identified for all selected mussels as in Khaitov et al. (2021) based on the presence (T-morphotype) or absence (E-morphotype) of an uninterrupted strip of prismatic layer under the ligament on the inner side of the shell. The T-morphotype is characteristic of *MT*, and the E-morphotype is characteristic of *ME*. The proportion of morphotypes was converted to the proportion of *MT* (*Ptros*) in each sample, in pooled samples from each substrate from each site (denoted as *Ptros*_{Algae} and *Ptros*_{Bottom}) and in pooled samples from each site (*Ptros*_{Site}), using the following equation:

$$P_{tros} = \frac{e^{-2.4 + 5.4PT}}{1 + e^{-2.4 + 5.4PT}}$$

where *PT* is the proportion of the T-morphotype.

This equation, derived from the 24 genotyped samples (in total, 1105 multilocus mussel genotypes) from the Kandalaksha Bay, accounts for the nearly linear dependence of *Ptros* on *PT* and reliably predicts *Ptros* over the entire salinity range in the White Sea, i.e. up to 24 psu (Khaitov et al. 2021). However, as studies in

the Barents Sea have shown, this equation may overestimate *Ptros* at higher salinities, e.g. up to 20‰ at salinity around 30 psu (Khaitov et al. 2021, Marchenko et al. 2023).

2.2.2. Assessment of environmental parameters

In total, we used 7 parameters describing possible influence of rivers, ports, surf and substrate on mussels (Table 1). We used 3 different proxies of salinity (RiverSize, DistRiver and Salinity) because a single estimate of salinity at low tide could be insufficient to characterize overall salinity and river influence per se at the sampling sites. Salinity was measured directly with an accuracy of 1 psu using an Atago S/Mill-E refractometer. To classify rivers by size (RiverSize), the data from Table S1 were used. To calculate Fetch, the R package 'waver' (Marchand & Gill 2018) was applied to regional shape-files.

2.3. Testing data sets

Three data sets were used for testing: 1 from the same area of the White Sea and 2 from the Barents Sea. The 'Kandalaksha littoral' data set contained 23 samples from 12 littoral sites in Kandalaksha Bay

Table 1. Environmental parameters involved in the study

Environmental parameter/ model predictor	Type	Explanation	Range (median) in the data
Influence of substrate			
Substrate	Categorical	Algal and Bottom samples for each site are treated separately	Algae vs. Bottom
Influence of rivers			
Salinity	Continuous	Surface salinity (psu) at the time of sampling, i.e. at low tide	2–30 (19)
DistRiver	Continuous	Straight-line distance (km) between the site and the nearest river mouth by map; the values were log-transformed when used for model fitting	0–18.5 (4.9)
RiverSize	Categorical	Rivers are categorized according to whether their catchment area is larger or smaller than the median area of all rivers in the region	Small vs. large
Influence of ports			
DistPort	Continuous	Straight-line distance (km) between the site and the nearest port by map; log-transformed values were used.	0.1–82.2 (18.7)
PortStatus	Categorical	Ports were categorized according to whether they are active or abandoned	Active vs. Abandoned
Influence of surf			
Fetch	Continuous	Unobstructed length of water surface (km) over which wind from a certain direction can blow; log-transformed values were used	0.2–28.8 (3.3)

(Fig. S1 in Supplement 2 at www.int-res.com/articles/suppl/m773pXXX_supp2.pdf). We took only algal samples at 4 sites, only bottom samples at 4 other sites and samples from both substrates at the remaining 4 sites (Table S4). Environmental parameters were assessed in the same way as for the modeling data set.

'Tyuva littoral' and 'Tyuva sublittoral' testing data sets were extracted from the published data of Marchenko et al. (2023). These authors mapped in detail the distribution of *Ptros* in mussel settlements of the 3 km long Tyuva Inlet in Kola Bay of the Barents Sea (Fig. 1) sampled in 2009–2010. *Ptros* was predicted either by direct genotyping or from morphotype frequencies using the formula derived for local populations existing under salinities higher than in the White Sea (Marchenko et al. 2023). They provided a number of environmental characteristics including depth, salinity, cover of macrophytes in rank scale and dominant algal species (usually, kelps in the sublittoral and fucoids in the littoral) for each sampling site. The 'Tyuva littoral' set contained samples from all 23 littoral sites from the depth range corresponding to the fucoid belt (0.5–1.5 m above mean spring tidal depth, Marchenko et al. 2023; note that the position of fucoid belt in the Barents Sea differs from that in the White Sea due to the different tidal amplitude). 'Tyuva sublittoral' contained samples from all 15 sublittoral sites (depth range from 0.5 to 3.5 m). Since the substrate of mussel fouling was not registered during sampling, we classified samples into bottom and algal substrate based on the algal cover at the sites (ranks 1–3 and 4–5, correspondingly). The remaining environmental parameters were assessed as for the modeling data set, with the nearest port in Ekaterininskaya Gavan Bight considered as active and the river Tyuva flowing into the inlet as a large one.

2.4. Statistical analysis

All data were processed using the statistical programming language R v.4.0.5 (R Core Team 2023).

2.4.1. Dependency of *Ptros* on environmental parameters in the modeling data set (Model 1)

We used a generalized additive mixed model (GAMM; Wood 2017) as a modeling technique, which has been shown to work well for SDM construction (Elith et al. 2006). One of the strengths of this approach is that additive models assume that the relationship between the dependent variable (in our case,

Ptros) and continuous predictors is not necessarily linear but may be curvilinear (Austin 2002). The weakness of the approach is that it does not provide a direct assessment of either relative or absolute importance of factors.

The fitted GAMM (hereafter, Model 1) was based on the distribution of beta-binomial residuals and on the restricted maximum likelihood method for estimation of the parameters. Smoothers for all continuous predictors were fitted using cubic basic splines. Categorical predictors were included as parametric terms in the model. Site was considered as a random factor. The function 'gam()' from the package 'mgcv' (Wood 2017) was used to fit the model.

To check for collinearity of all predictors in Model 1 and other models, we calculated the variance inflation factor (VIF; Fox & Monette 1992), considering values less than 3.5 as acceptable (Quinn & Keough 2002). To verify that Model 1 met the assumptions of sampling independence, we examined the presence of residuals' spatial autocorrelation by means of spline correlogram construction (Bjørnstad & Falck 2001) with the function 'spline.correlog()' from the package 'ncf' (Bornstad 2022) and found no evidence of spatial autocorrelation. We also considered the model residuals in relation to year of sampling and found no significant patterns.

2.4.2. Dependence of abundance of different morphotypes on environmental parameters in a modeling data set (Model 2)

The relationships between the taxonomic structure and the predictors were further investigated using abundances of mussels of different morphotypes. This means that we equated morphotypes with species. However, this assumption should not have crucially biased the results of the analysis, given the proportional relationship between *PT* and *Ptros* in mussel settlements in the study area (Khaitov et al. 2021). The mean abundances of mussels of E- and T-morphotypes across samples from each site (from both substrates) were log-transformed and used as dependent variables in Model 2. The model was constructed as a generalized additive model (GAM) with Gaussian distribution of residuals and included the factor Morphotype, the same set of predictors as in Model 1 except Substrate and Site, and interactions between Morphotype and other predictors. To confirm that Model 2 satisfies the assumptions of regression analysis, we inspected the residual plots. No discernible patterns were detected in the residuals.

2.4.3. Association between *Ptros*, substrate and mussel abundance

The ultimate goal of the analysis was to find out how the segregation of *ME* and *MT* between algal and bottom substrates depended on the abundance of each species on each substrate. For each site, we calculated the difference between the proportion of *MT* in algal and bottom samples: $\text{Diff} = \text{Ptros}_{\text{Algae}} - \text{Ptros}_{\text{Bottom}}$. The obtained Diff values were used as a dependent variable in Model 3, which was constructed as a GAM with Gaussian residual distribution.

Assessing the dependence of Diff on $\text{Ptros}_{\text{Site}}$, we could not directly operate with densities of morphotypes because they were collinear on different substrates (see Section 3). Therefore, we performed principal component analysis (PCA) for the abundance matrix of T- and E-morphotypes on algal and bottom substrates and used PC1 and PC2 values as proxies of morphotype abundances, along with $\text{Ptros}_{\text{Site}}$ in Model 3. As in previous cases, residual plots were examined to validate Model 3, and no violations of the regression analysis assumptions were detected.

2.4.4. Assessment of predictive power of Model 1

To check whether Model 1 could be used to predict the dominant species in bottom and algal samples at a site with known environmental parameters, *MT* ($\text{Ptros} > 0.5$) or *ME* ($\text{Ptros} < 0.5$), we used all of the parameters from Model 1 to predict $\text{Ptros}_{\text{Algae}}$ and $\text{Ptros}_{\text{Bottom}}$ for each site within the modeling data set and within each of the 3 testing data sets. The predicted values were categorized into those > 0.5 and those < 0.5 and considered to be classifiers for detecting *MT*- or *ME*-dominated samples. The receiver operating characteristics (ROC) followed by the analysis of the area under the curve (AUC; Fielding & Bell 1997, Fawcett 2006) were used to evaluate the performance of the models. We considered $\text{AUC} \geq 0.7$ as acceptable discrimination (Hosmer et al. 2013). The function 'roc()' from the package 'pROC' (Robin et al. 2011) was used.

3. RESULTS

Ranges and median values of the continuous predictors are summarized in Table 1. While the distribution of Fetch and Salinity values was highly variable, the most wind-exposed sites were located on the southeastern coast of the Bay and on open shores of the islands at its top (Fig. 1A) while the most desali-

nated areas were located at the very top of the Bay (Fig. 1C). Expectedly, Salinity tended to decrease towards river mouths (Fig. S2A in Supplement 2) and was lower closer to large rivers than to small ones (Fig. S2B). Sites close to ports tended to have lower Fetch (Fig. S2C), but no association between DistPort and Salinity was observed (Fig. S2D). All Pearson's correlations between Salinity, DistRiver, DistPort and Fetch were rather low (Table S5), the largest being the correlation between Fetch and DistPort ($r = 0.525$).

A striking feature visible on maps of *Ptros* distribution across algal and bottom substrates is the universally elevated proportion of *MT* on the former (Fig. 1D–G). While spatial distribution of *Ptros* was highly variable, its maximum values on both substrates were observed at the top of the Bay and in some deep inlets, while its minimum values were observed along the open part of the southeastern coast (Fig. 1D–G). It is difficult to discern relationships between *Ptros* and any environmental predictors other than substrate in the small-scale maps shown in Fig. 1. For this purpose, it is necessary to consider Model 1.

3.1. Relationship of *Ptros* and environmental parameters evaluated by Model 1

Although some non-zero pairwise correlations between environmental factors were found (see above), VIF values calculated for the predictors were generally low (maximal VIF being that for Fetch, 1.76). This result means that the collinearity between the predictors was negligible, i.e. they did not mask each other's influence.

Model 1 explained 77% of the total deviance. It revealed a significant dependency of *Ptros* on all predictors except DistRiver. Effective degrees of freedom (edf) for DistPort and Fetch were close to 1, indicating the linear dependence of *Ptros* on these predictors. In contrast, the dependence on the third continuous predictor, Salinity, was curvilinear (edf > 2 , Table 2).

According to the model, *Ptros* decreased both with DistPort (Fig. 2E) and with Fetch (Fig. 2G). This means that the proportion of *MT* was higher near ports and in surf-protected areas. PortStatus also had a significant effect: predicted *Ptros* was higher near active ports than near abandoned ones (Fig. 2I). The curvilinear dependence of *Ptros* on salinity can be described as follows: predicted *Ptros* decreases with salinity in the range from low to about 24 psu (average salinity in the White Sea) and increases again at higher salinities (up to 30 psu) (Fig. 2A). In addition,

Table 2. Parameters of smoothers and coefficients of parametric terms for Model 1 describing dependency of the proportion of *Mytilus trossulus* in mixed settlements (*Ptros*) on environmental predictors. edf: effective degrees of freedom; ref.edf: reference effective degrees of freedom

Smoother terms	edf	ref.edf	Chi squared	p
s(Salinity)	2.4	9	396.7	0.003
s(DistRiver)	0.0	9	0.0	0.672
s(Fetch)	0.9	9	88.2	0.042
s(DistPort)	1.0	9	276.2	0.002
Random effect s(Site)	74.4	92	453.6	<0.0001
Parametric terms	Parameter estimate	SE	z-statistic	p
(Intercept)	-1.7	0.1	-11.8	<0.0001
Substrate _(Algae)	0.9	0.1	14.6	<0.0001
RiverSize _(Large)	0.4	0.2	2.6	0.009
PortStatus _(Active)	1.0	0.2	5.7	<0.0001

predicted *Ptros* was higher near large rivers than near small ones (Fig. 2I). Finally, *Ptros* was higher on algal substrates than on bottom ones (Fig. 2I, see also Fig. 1D,F).

3.2. Dependence of abundance of different morphotypes on environmental parameters evaluated by Model 2

The results of Model 2 were in good agreement with those of Model 1 for all predictors, i.e. Salinity (Fig. 2B), RiverSize (Fig. 2J), DistPort (Fig. 2F), PortStatus (Fig. 2J), and Fetch (Fig. 2H) (see Table S6 for all model parameters). In addition, they revealed an asymmetry in the responses of the 2 species to some of these predictors. While the abundance of T-morphotypes did not vary with Salinity, that of E-morphotypes dropped at low salinity (Fig. 2B). On the other hand, the abundance of E-morphotypes slightly var-

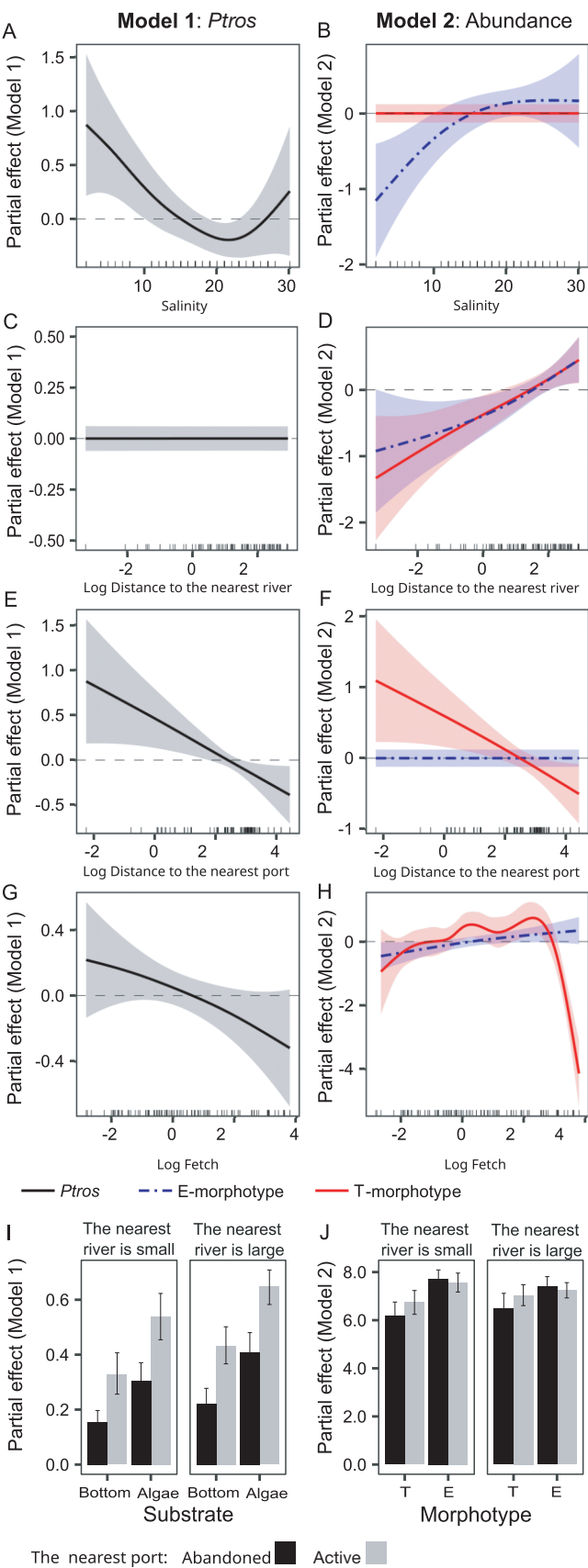


Fig. 2. Partial effects of environmental parameters on either proportion of *Mytilus trossulus* in mixed settlements (*Ptros*) or abundance of species-specific morphotypes evaluated by fitted generalized additive (mixed) models (Model 1 and Model 2, respectively). Dependency of *Ptros* on (A) salinity at low tide (Salinity), (C) distance to the nearest river (DistRiver), (E) distance to the nearest port (DistPort), (G) wind exposure (Fetch), and (I) discrete predictors: substrate type, nearest port status and size of the nearest river. Dependency of E- and T-morphotype log-transformed abundances on (B) Salinity, (D) DistRiver, (F) DistPort, (H) Fetch, and (J) nearest port status and size of the nearest river. Ribbons around curves and whiskers represent 95% confidence intervals

ied with Fetch and DistPort, while that of T-morphotypes strongly decreased both with the distance from ports (Fig. 2F) and with surf level (Fig. 2H).

3.3. Dependency of *Ptros* on substrate and mussel abundance evaluated by Model 3

The PCA of the abundance matrix of T- and E- morphotypes revealed a high positive correlation of PC1 (explained 62% of total variation) with abundances of T-morphotypes and of PC2 (20% of total variation) with abundances of E-morphotypes on both substrates (Fig. 3B,C). Thus, the abundance of conspecific morphotypes varied consistently on different substrates (see also Fig. 1D,F). Therefore, PC1 and PC2 can be considered as proxies of *MT* and *ME* abundance, respectively.

Parameters of Model 3, which explained 31% of the deviance, are provided in Table S7. Fig. 3 demonstrates how the difference between *MT* proportion on algal and bottom substrates (Diff) depends on *MT* prevalence at the site (*Ptros*_{Site}) and mussel abundances in terms of PCs according to the model. The dependence of Diff on *Ptros*_{Site} was significant ($p < 0.001$) and bell-shaped, with minimal values at sites absolutely dominated by *ME* or *MT* (*Ptros* close to 0 or 1) and maximal at sites with equal presence of both species (Fig. 3A). This means that the strongest segregation of the 2 species by substrate was observed in settlements where they occurred in equal proportions. Dependence of Diff on PC1 was marginally significant ($p = 0.087$) and tended to decrease with increasing PC1 (Fig. 3B). The dependence of Diff on PC2 was significantly positive ($p = 0.011$, Fig. 3C; Table S6). This means that the species were strongly

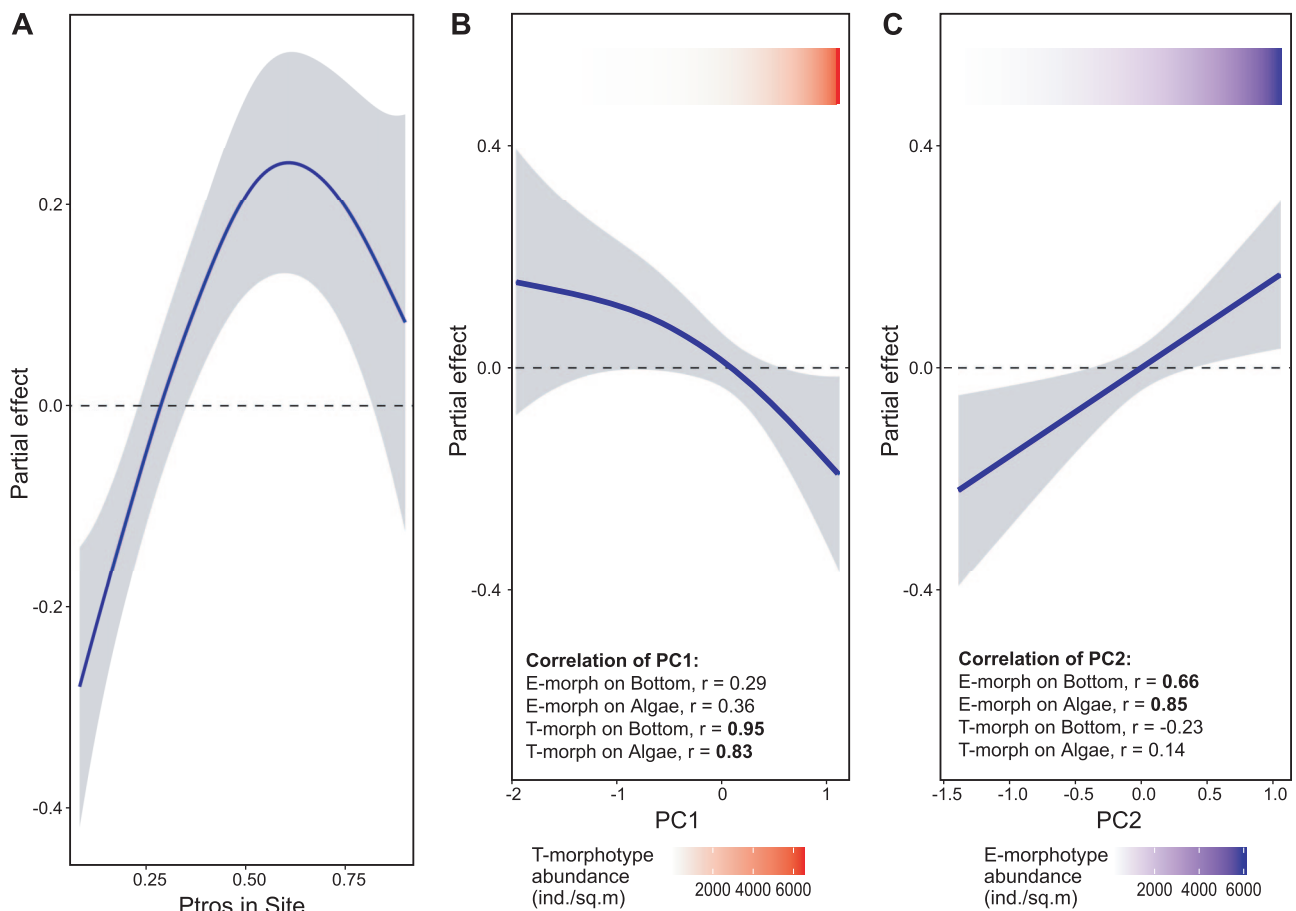


Fig. 3. (A) Dependence of the difference between proportion of *Mytilus trossulus* on algal and bottom substrates (Diff) on the proportion of *M. trossulus* at a site (*Ptros*_{Site}), and estimations of total abundance of (B) *M. trossulus* and (C) *M. edulis*. Principal components from the matrix of T- and E-morphotype abundances on different substrates are considered as proxies for *M. trossulus* and *M. edulis* abundances (PC1 and PC2, respectively). Correlations of mussel abundances with PC scores are given, with the 2 highest coefficients in **bold**. Ribbons around curves represent 95% confidence intervals. Colored gradient bars at the top inside panels (B) and (C) reflect linear associations between PC1 and T-morphotype and PC2 and E-morphotype abundance, respectively

segregated by substrates at sites with a high *ME* abundance but not at sites with a high *MT* abundance (PC1 and PC2 are proxies for *MT* and *ME* abundances, respectively).

3.4. Assessment of the predictive power of Model 1

Model 1 fit well for the 'Kandalaksha littoral' testing data set (AUC = 0.85 vs. AUC = 0.84 for the modeling data set). It classified the samples into *ME*- and *MT*-dominated sites fairly well, with the exception of a few false negatives, i.e. sites unpredictably dominated by *MT* (Fig. 4A,B). Its predictive power for the 2 testing sets from the Barents Sea was lower but

acceptable: AUC = 0.71 for 'Tyuva littoral' and AUC = 0.69 for 'Tyuva sublittoral' (Fig. 4C,D). In contrast to the 'Kandalaksha littoral' testing data set, most of the false results were positive.



4. DISCUSSION

Almost all environmental predictors considered in our study — namely, surf level, distance to the port, status of the port, salinity at low tide, size of the nearest river and fouling substrate — influenced the distribution of *ME* and *MT* in the White Sea. The differences in the distribution reflected the partial divergence of ecological niches of these 2 species.

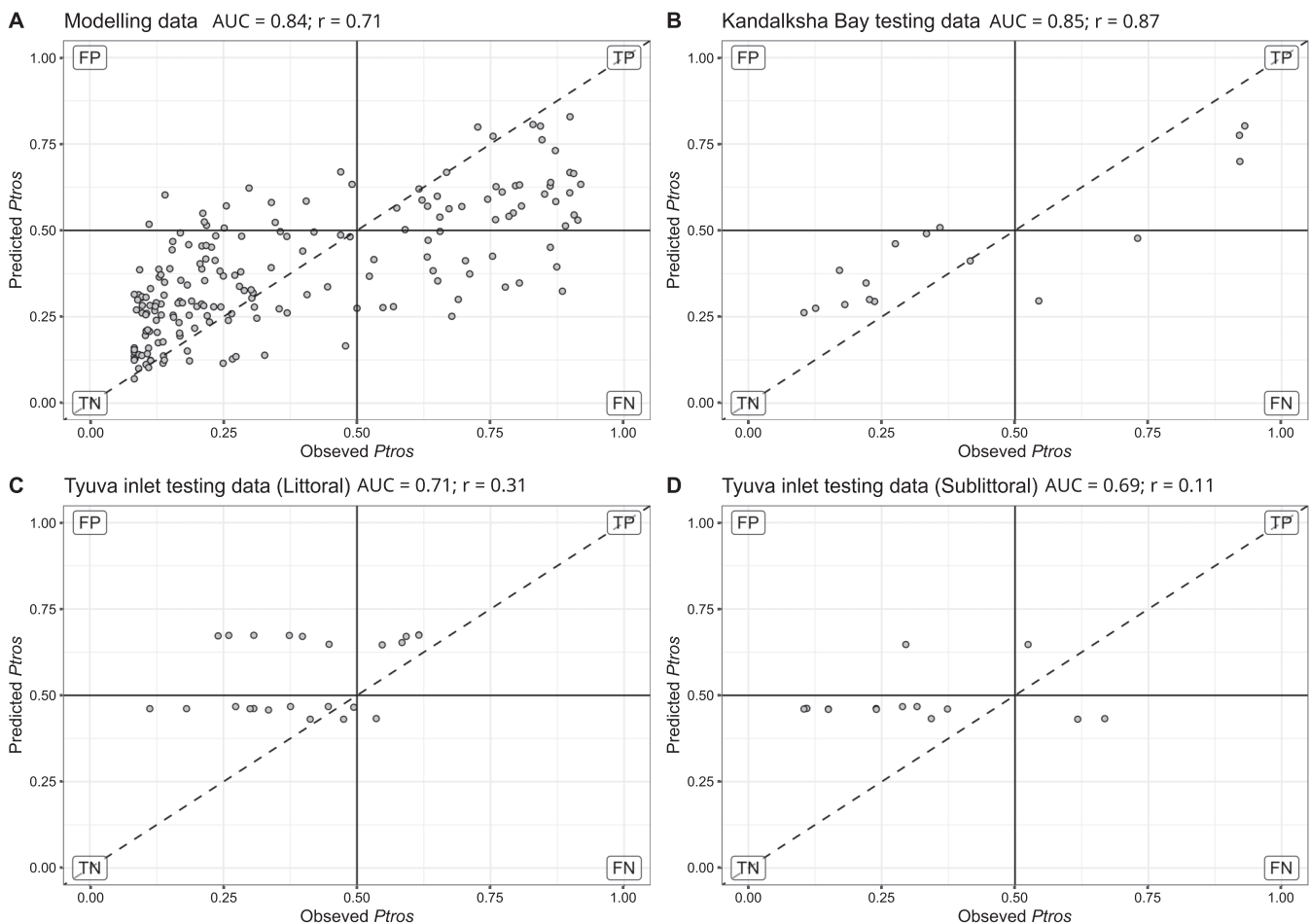


Fig. 4. Ability of the species distribution model (Model 1) to predict the proportion of *Mytilus trossulus* (*Ptros*) in mussel samples from the (A) modeling and the (B–D) testing data sets. Each plot compares empirical *Ptros* in samples from algal and bottom substrates and *Ptros* predicted by the model within the particular data set. If the empirical and the predicted values were the same, the points would lie on the diagonal (dashed line). Labels mark the quadrants with false positive (FP), true positive (TP), true negative (TN) and false negative (FN) predictions in the analysis of the ability of the model to classify samples into *Mytilus edulis*-dominated (*Ptros* < 0.5) and *M. trossulus*-dominated (*Ptros* > 0.5) ones. Data set names are shown in panel headings. Values of the area under the receiver operating characteristic curve (AUC) for binary classification (*M. edulis*-dominated vs. *M. trossulus*-dominated) and Pearson correlations between observed and predicted *Ptros* are given in panel headings

4.1. Ecological niche partitioning between *MT* and *ME* in the Kola contact zone

We showed that the most expected habitat of *ME* in the White Sea littoral was a bottom substrate in a surf-exposed location with a surface salinity of 24 psu (average for the White Sea) situated away from ports and large rivers. The most expected habitat of *MT* was an algal substrate in a surf-protected location with a salinity lower than the White Sea average situated close to active ports and large rivers. While differences associated with substrate and distance to ports have been previously noted in the White Sea (Väinölä & Strelkov 2011, Katolikova et al. 2016), those associated with salinity and surf exposure were first uncovered in this study.

4.1.1. Segregation by salinity

In the Baltic Sea, *MT* is adapted to an extremely low salinity (Knöbel et al. 2021, Wiesenthal et al. 2025). Comparative ecophysiological data on *MT* and *ME* elsewhere are scarce (Gardner & Thompson 2001, Qiu et al. 2002, Sokolova et al. 2024). Before our study, there was no convincing evidence of segregation of these species by salinity in contact zones outside the Baltic (Moreau et al. 2005, Riginos & Cunningham 2005, Katolikova et al. 2016, Marchenko et al. 2023). For the White Sea, this lack of evidence could be due to at least 3 reasons. Firstly, the role of salinity may be masked by other factors such as distance to ports. This seems particularly plausible in light of our data that *MT* in the White Sea is a euryhaline species forming mass settlements in the entire salinity range recorded in our study, while *ME* is much less abundant at salinities below 15 psu (Fig. 2B). Secondly, in contrast to the Baltic, there are no broad geographic salinity gradients in the White Sea, only local ones.

The third reason is the curvilinear dependence of the proportion of *MT* in mixed settlements (*Ptros*) on salinity: *Ptros* increases not only when the salinity is reduced but also when it is extremely high for the White Sea (up to 30 psu, Fig. 2B). This nonlinearity, which may prevent the dependence from being detected, can be explained in 2 ways. On the one hand, local summer surface salinity above 24 psu in the Kandalaksha Bay, supposedly associated with irregular episodes of upwelling (Dale & Prego 2003), may be a nonspecific stress for littoral animals adapted to lower salinity. *MT* may be more tolerant to this stress. On the other hand, as shown in detailed studies in the Barents Sea (Khaitov et al. 2021, Marchenko et al.

2023), the method of predicting *Ptros* ('morphotype test') used in our study may overestimate it at salinities close to 30 psu. Therefore, we cannot rule out the possibility that the increased *Ptros* at sites with a high salinity is an artifact.

4.1.2. Non-random distribution depending on distance to ports

The association of *MT* with harbors in the White and the Barents Seas may be associated with its invasion with maritime transport from the western Atlantic in the 20th century (Väinölä & Strelkov 2011). It has also been hypothesized that *MT* is more resistant to anthropogenic pollution and is better adapted to disturbed habitats than *ME* (Katolikova et al. 2016). Our observation that *MT* frequency is lower near abandoned ports than near active ones is consistent with this hypothesis. However, the propagule pressure of *MT* may have decreased near abandoned harbors in recent decades, which could affect the size of its populations.

4.1.3. Segregation by surf level

Segregation of *ME* and *MT* by surf level (*ME* is more common in localities exposed to surf, while *MT* is more common in sheltered areas) may be due to the well-known differences in the mechanical properties of their shells and the ability to form dense aggregations. *ME* has thicker, heavier and less flexible shells (Beaumont et al. 2008, Michalek et al. 2021) and is more inclined to form tight clumps (Liu et al. 2011). These features may be adaptive on exposed coasts. Theoretically, the distribution of *ME* and *MT* by surf level as well as substrate may also be affected by differences in byssus secretion and attachment strength. Unfortunately, there are no comparative data on this topic.

4.1.4. Segregation by substrate

Segregation by substrate (*ME* is more common on bottom substrates, while *MT* is more common on algal ones) may be explained by the same differences as segregation by surf level. Other things being equal, *MT*, with its thin fragile shells, should be lighter than *ME* (Michalek et al. 2021) and thus better suited to life on algae. In addition, fucoid thalli may serve as shock absorbers (Katolikova et al. 2016) and provide shelter

from starfish selectively preying on *MT* in mixed settlements (Khaitov et al. 2018, 2023). By the same token, denser aggregations formed by *ME* are more adaptive on the bottom than on algae.

4.1.5. Competition for substrate

Whatever historical, physiological, morphological, behavioral and other features influence the segregation of *MT* and *ME*, interspecific competition may also be involved. Assessing the role of mussel abundance in the segregation across substrates, we found that *MT* abundance did not significantly affect it but *ME* abundance did: as the latter increased, the degree of segregation increased as well (Fig. 3B,C). In our opinion, this pattern results from the divergence of the realized (but not the fundamental) species niches: *ME* outcompetes *MT* on bottom substrates, displacing it to algal thalli.

Spatial segregation of sympatric mussels by substrate, which is apparently density-dependent, is evident at the level of tens of centimeters (Katolikova et al. 2016). Segregation at a similarly small scale has been found in terrestrial plants (Raventós et al. 2010). A biologically generated spatial pattern model, which has been developed for plants, relates inter-specific segregation with the intra-specific clustering in competing species (Pacala & Levin 1997, Amarasekare 2003). Our findings suggest that this model can also be applied to mussels, which share with plants the similarity of being attached organisms.

4.2. Predictive power of SDM

The ability of our model to classify sites into *ME*- and *MT*-dominated ones in an independent testing data set from the White Sea was high (AUC = 0.85). Therefore, we assume that the predictors included in the model explain most of the variation in species distribution within the littoral fucoid belt in the White Sea.

The predictive performance of the model for the Barents Sea data (AUC \approx 0.7), while formally acceptable, was lower than for the White Sea. This may be due to the following reasons. Firstly, it may be associated with a large depth range of the sampling sites, considering that the distribution of *ME* and *MT* in the Tyuva Inlet by depth is non-random (Marchenko et al. 2023). The second reason may be a coarser categorization of the Barents Sea samples into algal and bottom ones. Since fouling substrate was not taken into

account during sampling, we predicted it based on the projective cover of algae at the sampling site. Thirdly, we do not know whether the 2 species are non-randomly distributed across bottom and algal substrates in the sublittoral, where fucoids are replaced by kelps. The fourth reason could be a narrow variation in the values of DistPort, DistRiver and Fetch in the small Tyuva Inlet in comparison with the Kandalaksha Bay.

Finally, the fact that SDM tended to overestimate *Ptros* in the Barents Sea data (false positive predictions) is consistent with the observation that the proportion of *MT* has been declining in the study area in the 2010s under seemingly stable environmental conditions in terms of predictors included in our model (Marchenko et al. 2023). This observation suggests the presence of some yet unstudied factors regulating the taxonomic structure.

4.3. Ecological niche partitioning between *MT* and *ME* in the Kola contact zone as compared to other zones

Blue mussels are a challenging model for studying ecological niche partitioning between cryptic species in sympatry due to their wide distribution, biogeographic history and hybridization. *ME* and *MT* play similar ecological roles in their native oceans, i.e. Atlantic and Pacific, respectively (compare Commito & Dankers 2001 and Bodkin et al. 2018) and may inherently have overlapping fundamental ecological niches. Contact zones between these species in the Atlantic can be considered as ecological (and evolutionary) experiments, set in strikingly different environments (from the Baltic to Spitsbergen) at different times (from post-glacial to the historical period, Väinölä & Strelkov 2011, Wenne et al. 2020 and references therein). The design of these experiments was possibly different too, because in some zones the original settler could be *ME* and in others, *MT*. In addition, competition (ecological character displacement, Pfennig & Pfennig 2009), hybridization (reinforcement of prezygotic reproductive isolation, Lukhtanov 2011) and introgression (adaptive introgression, Hedrick 2013) could influence the divergence of their ecological phenotypes differently in different zones. These considerations suggest that the zones should differ, and this hypothesis has been a recurrent theme in genetic research on blue mussel contact zones (Riginos & Cunningham 2005, Bierne et al. 2016, Fraïsse et al. 2016). Nevertheless, we believe that the differences between *ME* and *MT* are more fundamental and that

conspecific ecological phenotypes in different zones should thus be similar, producing comparable patterns in species distributions. Some results of this study support this hypothesis.

The observation that *MT* frequency is elevated in low-salinity habitats not only in the Baltic but also in the White Sea seems to resolve the old conundrum about seemingly contrasting salinity adaptations of the Baltic and other Atlantic *MT* populations (e.g. Riginos & Cunningham 2005, Katolikova et al. 2016, [see also above](#)). Further, an increased *MT* frequency has been repeatedly observed in calm and freshened waters, e.g. in the tops of fjords near Bergen in Norway (Ridgway & Nævdal 2004) and Uummannaq in Greenland (Wenne et al. 2020), and in Loch Etive in Scotland (Beaumont et al. 2008), which is hardly a coincidence. Our observations indicate that this combination of weak surf and low salinity is also favorable for *MT* in the White Sea.

With the exception of salinity in the Baltic, none of the predictors affecting segregation of *ME* and *MT* in the White Sea have been convincingly shown to act in other contact zones. Data on surf are inconsistent (compare Bates & Innes 1995, Comesaña et al. 1999, Tam & Scrosati 2014 and this study), and data on substrates are absent. If different preferences of *ME* and *MT* for surf level and substrates are indeed associated with their morphological and behavioral differences (Katolikova et al. 2016; this study), these preferences should be omnipresent. Ad hoc studies in other contact zones might shed light on this matter. It would also be interesting to examine the differences between *ME* and *MT* in tolerance to stress, particularly anthropogenic pollution, e.g. in harbors (see discussion in Brooks et al. 2015 and Beyer et al. 2017).

4.4. Strengths and weaknesses of our approaches to the study of sympatric mussels

The methods used in our study have certain limitations. We identified the mussels with the help of the 'morphotype test', which does not permit a direct assessment of species abundances and does not identify hybrids as a separate category. The latter limitation is not particularly significant in the Kola zone, where hybrids are relatively scarce (Väinölä & Strelkov 2011, Wenne et al. 2020), but may compromise studies in other contact zones, where hybrids may play an important ecological role (e.g. Schwartz et al. 2024).

Although *ME* and *MT* differ everywhere in morphotype frequencies, the magnitude of the differences varies between contact zones and between habitats

with different salinities in the Arctic (Khaitov et al. 2021). This means that the 'morphotype test' must be calibrated before use in a new area (see Khaitov et al. 2021 for recommendations). Multilocus genotyping, while still too costly for processing dozens of thousands of specimens needed for SDM, remains the gold standard of taxonomic assessment in blue mussels.

We did not account for some potential predictors affecting species segregation such as depth (Marchenko et al. 2023), predators (Khaitov et al. 2018, 2023) or temperature (Kijewski et al. 2019). Moreover, some of our predictors could have been estimated more carefully (for example, bottom salinity at high tide could be more informative for littoral mussels than salinity at low tide). However, since most of our predictors were shown to be significant, they should not be ignored in future studies.

The correlative approach used in our study does not allow a direct assessment of either relative or absolute importance of factors. For instance, we cannot say whether salinity or substrate is more crucial. However, the take-home message from our research is that there is no single 'leading' factor determining the distribution of *ME* and *MT*, contrary to the idea that has dominated the field since the pioneering studies in the Baltic (e.g. Gardner & Thompson 2001, Ridgway & Nævdal 2004, Riginos & Cunningham 2005, Śmietanka et al. 2014).

The limitations discussed above do not detract from the fact that, as shown in our study, SDMs may be a useful tool to study the distributions of *ME* and *MT* in sympatry. Their obvious benefits include the possibility to analyze the distributions of the species in the space of multiple predictors simultaneously, the possibility to control the collinearity of the predictors and the lack of necessity to treat dependencies as linear.

Promising directions of further research on niche partitioning in sympatric mussel species are as follows. The use of taxonomic methods allowing direct assessment of abundances of species and their hybrids would elucidate the nature of competition between them all. Incorporation of additional environmental factors into SDMs might yield surprising results. Further, it would be worthwhile to take a closer look at different spatial scales, down to the smallest one, in the segregation of these 2 mussel species. Finally, a parallel study in different contact zones would reveal common and zone-specific patterns. The classical review on the divergence of ecological niches of *ME* and *MT* in different contact zones (Riginos & Cunningham 2005) is 20 yr old. The time is ripe for a new survey, and our observations from the Kola zone may prove useful.



Acknowledgements. We thank all participants of our expeditions, especially Gita Paskerova and Eugene Genelt-Yanovsky, for their help in fieldwork. We are deeply grateful to Natalia Lentsman for the English language assistance during the preparation of the manuscript. We express our special gratitude to the administration of the Kandalaksha State Nature Reserve, where a significant part of the material was collected, for comprehensive assistance. We also thank 3 anonymous reviewers for their constructive comments that made it possible to improve the paper considerably. This study was supported by the Russian Science Foundation, grant no. 19-74-20024.

LITERATURE CITED

- Amarasekare P (2003) Competitive coexistence in spatially structured environments: a synthesis. *Ecol Lett* 6: 1109–1122
- Austin MP (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecol Model* 157:101–118
- Bates J, Innes D (1995) Genetic variation among populations of *Mytilus* spp. in eastern Newfoundland. *Mar Biol* 124: 417–424
- Beaumont AR, Hawkins MP, Doig FL, Davies IM, Snow M (2008) Three species of *Mytilus* and their hybrids identified in a Scottish Loch: natives, relicts and invaders? *J Exp Mar Biol Ecol* 367:100–110
- Berger VY, Naumov A (2000) General features of the White Sea. *Ber Polarforsch* 359:3–9
- Berger V, Dahle S, Galaktionov K, Kosobokova X and others (2001) White Sea. Ecology and environment. Derzavets Publishers, St-Petersburg-Tromsø
- Beyer J, Green NW, Brooks S, Allan IJ and others (2017) Blue mussels (*Mytilus edulis* spp.) as sentinel organisms in coastal pollution monitoring: a review. *Mar Environ Res* 130:338–365
- Bickford D, Lohman DJ, Sodhi NS, Ng PKL and others (2007) Cryptic species as a window on diversity and conservation. *Trends Ecol Evol* 22:148–155
- Bierne N, Welch J, Loire E, Bonhomme F, David P (2011) The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Mol Ecol* 20:2044–2072
- Bjørnstad ON (2022) ncf: spatial covariance functions. **R package version 1.3-2.**
- Bjørnstad ON, Falck W (2001) Nonparametric spatial covariance functions: estimation and testing. *Environ Ecol Stat* 8:53–70
- Bodkin JL, Coletti HA, Ballachey BE, Monson DH, Esler D, Dean TA (2018) Variation in abundance of Pacific blue mussel (*Mytilus trossulus*) in the Northern Gulf of Alaska, 2006–2015. *Deep Sea Res II* 147:87–97
- Brooks SJ, Farnen E, Heier LS, Blanco-Rayón E, Izagirre U (2015) Differences in copper bioaccumulation and biological responses in three *Mytilus* species. *Aquat Toxicol* 160:1–12
- Buschbaum C, Dittmann S, Hong JS, Hwang IS and others (2009) Mytilid mussels: global habitat engineers in coastal sediments. *Helgol Mar Res* 63:47–58
- Caradima B, Schuwirth N, Reichert P (2019) From individual to joint species distribution models: a comparison of model complexity and predictive performance. *J Biogeogr* 46:2260–2274
- Comesaña AS, Toro JE, Innes DJ, Thompson RJ (1999) A molecular approach to the ecology of a mussel (*Mytilus edulis* – *M. trossulus*) hybrid zone on the east coast of Newfoundland, Canada. *Mar Biol* 133:213–221
- Commuto JA, Dankers NMJA (2001) Dynamics of spatial and temporal complexity in European and North American soft-bottom mussel beds. In: **Reise K (ed) Ecological comparisons of sedimentary shores. Ecological Studies, Vol 151. Springer, Berlin, p 39–59**
- Dale AW, Prego R (2003) Tidal and seasonal nutrient dynamics and budget of the Chupa Estuary, White Sea (Russia). *Estuar Coast Shelf Sci* 56:377–389
- Dennis AB, Hellberg ME (2010) Ecological partitioning among parapatric cryptic species. *Mol Ecol* 19:3206–3225
- Druehl LD, Green JM (1982) Vertical distribution of intertidal seaweeds as related to patterns of submersion and emersion. *Mar Ecol Prog Ser* 9:163–170
- Dufresnes C, Poyarkov N, Jablonski D (2023) Acknowledging more biodiversity without more species. *Proc Natl Acad Sci USA* 120:e2302424120
- Elith J, Leathwick JR (2009) Species distribution models: ecological explanation and prediction across space and time. *Annu Rev Ecol Evol Syst* 40:677–697
- Elith J, Graham CH, Anderson RP, Dudík M and others (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129–151
- Fawcett T (2006) An introduction to ROC analysis. *Pattern Recognit Lett* 27:861–874
- Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ Conserv* 24:38–49
- Filatov N, Pozdnyakov D, Johannessen OM, Pettersson LH, Bobylev LP (2007) White Sea: its marine environment and ecosystem dynamics influenced by global change. **Springer Science & Business Media, XXXXXX**
- Fox J, Monette G (1992) Generalized collinearity diagnostics. *J Am Stat Assoc* 87:178–183
- Fraïsse C, Belkhir K, Welch JJ, Bierne N (2016) Local inter-species introgression is the main cause of extreme levels of intraspecific differentiation in mussels. *Mol Ecol* 25: 269–286
- Gardner JPA, Thompson RJ (2001) The effects of coastal and estuarine conditions on the physiology and survivorship of the mussels *Mytilus edulis*, *M. trossulus* and their hybrids. *J Exp Mar Biol Ecol* 265:119–140
- Gardner JPA, Oyarzun PA, Toro JE, Wenne R, Zbawicka M (2021) Phylogeography of Southern Hemisphere blue mussels of the genus *Mytilus*: evolution, biosecurity, aquaculture and food labelling. **Oceanogr Mar Biol Annu Rev** 59:139–228
- Geller JB, Darling JA, Carlton JT (2010) Genetic perspectives on marine biological invasions. *Annu Rev Mar Sci* 2: 367–393
- Gosling E (2021) Marine mussels: ecology, physiology, genetics and culture. **John Wiley & Sons, XXXXXX**
- Hayhurst S, Rawson PD (2009) Species-specific variation in larval survival and patterns of distribution for the blue mussels *Mytilus edulis* and *Mytilus trossulus* in the Gulf of Maine. *J Molluscan Stud* 75:215–222
- Hedrick PW (2013) Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Mol Ecol* 22: 4606–4618
- Hellou J, Law RJ (2003) Stress on stress response of wild mussels, *Mytilus edulis* and *Mytilus trossulus*, as an indicator of ecosystem health. *Environ Pollut* 126:407–416

- Hosmer DW Jr, Lemeshow S, Sturdivant RX (2013) Applied logistic regression, 3rd edn. **John Wiley & Sons**.
- Hu ZM, Zhang QS, Zhang J, Kass JM and others (2021) Intraspecific genetic variation matters when predicting seagrass distribution under climate change. *Mol Ecol* 30: 3840–3855
- Katolikova M, Khaitov V, Väinölä R, Gantsevich M, Strelkov P (2016) Genetic, ecological and morphological distinctness of the blue mussels *Mytilus trossulus* Gould and *M. edulis* L. in the White Sea. *PLOS ONE* 11:e0152963
- Khaitov V, Makarycheva A, Gantsevich M, Lentsman N and others (2018) Discriminating eaters: Sea stars *Asterias rubens* L. feed preferably on *Mytilus trossulus* Gould in mixed stocks of *Mytilus trossulus* and *Mytilus edulis* L. *Biol Bull* 234:85–95
- Khaitov V, Marchenko J, Katolikova M, Väinölä R and others (2021) Species identification based on a semi-diagnostic marker: evaluation of a simple conchological test for distinguishing blue mussels *Mytilus edulis* L. and *M. trossulus* Gould. *PLOS ONE* 16:e0249587
- Khaitov VM, Makarycheva AY, Nematova RB, Evdokimova AI (2023) Predators regulate the taxonomic structure of mixed *Mytilus edulis* L. and *M. trossulus* Gould settlements in the shallow waters of the White Sea. *Proc Zool Inst Russ Acad Sci* 327:8–24
- Kijewski T, Zbawicka M, Strand J, Kautsky H, Kotta J, Rätsep M, Wenne R (2019) Random forest assessment of correlation between environmental factors and genetic differentiation of populations: case of marine mussels *Mytilus*. *Oceanologia* 61:131–142
- Knöbel L, Nascimento-Schulze JC, Sanders T, Zeus D and others (2021) Salinity driven selection and local adaptation in Baltic Sea mytilid mussels. *Front Mar Sci* 8: 692078
- Knowlton N (1993) Sibling species in the sea. *Annu Rev Ecol Syst* 24:189–216
- Koehn RK (1991) The genetics and taxonomy of species in the genus *Mytilus*. *Aquaculture* 94:125–145
- Krasavtsev LB (2011) Russia's foreign trade through the ports of the White Sea in the early twentieth century. *Vestnik Severnogo (Arkticheskogo) federal'nogo universiteta Seriya: Gumanitarnye i social'nye nauki* 1:17–24 (in Russian)
- Lindgren M, Gabellini AP, Munk P, Edelvang K, Hansen FT (2022) Identifying key processes and drivers affecting the presence of non-indigenous marine species in coastal waters. *Biol Invasions* 24:2835–2850
- Liu G, Stapleton E, Innes D, Thompson R (2011) Aggregational behavior of the blue mussels *Mytilus edulis* and *Mytilus trossulus*: a potential pre-zygotic reproductive isolation mechanism. *Mar Ecol* 32:480–487
- Lowen JB, Hart DR, Stanley RRE, Lehnert SJ, Bradbury IR, Dibacco C (2019) Assessing effects of genetic, environmental, and biotic gradients in species distribution modelling. *ICES J Mar Sci* 76:1762–1775
- Lukhtanov V (2011) Dobzhansky's rule and reinforcement of prezygotic reproductive isolation in zones of secondary contact. *Biol Bull Rev* 1:2–12
- Marchand P, Gill D (2018) waver: calculate fetch and wave energy. **R package version 0.2.1**.
- Marchenko J, Khaitov V, Katolikova M, Sabirov M and others (2023) Patterns of spatial and temporal dynamics of mixed *Mytilus edulis* and *M. trossulus* populations in a small subarctic inlet (Tyuva Inlet, Barents Sea). *Front Mar Sci* 10:1146527
- Michalek K, Vendrami DLJ, Bekaert M, Green DH and others (2021) *Mytilus trossulus* introgression and consequences for shell traits in longline cultivated mussels. *Evol Appl* 14:1830–1843
- Morais P, Reichard M (2018) Cryptic invasions: a review. *Sci Total Environ* 613-614:1438–1448
- Moreau V, Tremblay R, Bourget E (2005) Distribution of *Mytilus edulis* and *M. trossulus* on the Gaspé coast in relation to spatial scale. *J Shellfish Res* 24:545–551
- Ovaskainen O, Abrego N (2020) Joint distribution modelling (with applications in R). **Cambridge University Press**.
- Pacala SW, Levin SA (1997) Biologically generated spatial pattern and the coexistence of competing species. In: Tilman D, Kareiva P (eds) *Spatial ecology: the role of space in population dynamics and interspecific interactions*. Monographs in population biology 30. Princeton University Press, Princeton, NJ, p 204–232
- Penney RW, Hart MJ, Templeman N (2002) Comparative growth of cultured blue mussels, *Mytilus edulis*, *M. trossulus* and their hybrids, in naturally occurring mixed-species stocks. *Aquacult Res* 33:693–702
- Peterson ML, Doak DF, Morris WF (2019) Incorporating local adaptation into forecasts of species' distribution and abundance under climate change. *Glob Change Biol* 25:775–793
- Pfennig KS, Pfennig DW (2009) Character displacement: ecological and reproductive responses to a common evolutionary problem. *Q Rev Biol* 84:253–276
- Poggiato G, Münkemüller T, Bystrova D, Arbel J, Clark JS, Thuiller W (2021) On the interpretations of joint modeling in community ecology. *Trends Ecol Evol* 36: 391–401
- Qiu JW, Tremblay R, Bourget E (2002) Ontogenetic changes in hyposaline tolerance in the mussels *Mytilus edulis* and *M. trossulus*: implications for distribution. *Mar Ecol Prog Ser* 228:143–152
- Quinn GP, Keough MJ (2002) *Experimental design and data analysis for biologists*. **Cambridge University Press**.
- R Core Team (2023) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Raventós J, Wiegand T, Luis MD (2010) Evidence for the spatial segregation hypothesis: a test with nine-year survivorship data in a Mediterranean shrubland. *Ecology* 91: 2110–2120
- Reiss H, Cunze H, König K, Neumann K, Kröncke I (2011) Species distribution modelling of marine benthos: a North Sea case study. *Mar Ecol Prog Ser* 442:71–86
- Ridgway G, Nævdal G (2004) Genotypes of *Mytilus* from waters of different salinity around Bergen, Norway. *Helgol Mar Res* 58:104–109
- Riginos C, Cunningham CW (2005) Local adaptation and species segregation in two mussel (*Mytilus edulis* x *Mytilus trossulus*) hybrid zones. *Mol Ecol* 14:381–400
- Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Müller M (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 12:77
- Sáez AG, Lozano E (2005) Body doubles. *Nature* 433:111
- Sailing directions of the White Sea (1932) Leningrad: izdanie i tipografiya Gidrograficheskogo upravleniya.**
- Schwartz LC, González VL, Strong EE, Truebano M, Hilbish TJ (2024) Transgressive gene expression and expression plasticity under thermal stress in a stable hybrid zone. *Mol Ecol* 33:e17333

- Simon A**, Fraïsse C, El Ayari T, Liautard-Haag C, Strelkov P, Welch JJ, Bierne N (2021) How do species barriers decay? Concordance and local introgression in mosaic hybrid zones of mussels. *J Evol Biol* 34:208–223
- Śmietanka B, Burzyński A, Hummel H, Wenne R (2014) Glacial history of the European marine mussels *Mytilus*, inferred from distribution of mitochondrial DNA lineages. *Heredity* 113:250–258
- Sokolova IM, Kovalev A, Timm S, Marchenko J, Sukhotin A (2024) Species-specific metabolome changes during salinity downshift in sub-Arctic populations of *Mytilus edulis* and *M. trossulus*. *Front Mar Sci* 11:1403774
- Struck TH, Feder JL, Bendiksy M, Birkeland S and others (2018) Finding evolutionary processes hidden in cryptic species. *Trends Ecol Evol* 33:153–163
- Stuckas H, Knöbel L, Schade H, Breusing C and others (2017) Combining hydrodynamic modelling with genetics: Can passive larval drift shape the genetic structure of Baltic *Mytilus* populations? *Mol Ecol* 26:2765–2782
- Tam JC, Scrosati RA (2014) Distribution of cryptic mussel species (*Mytilus edulis* and *M. trossulus*) along wave exposure gradients on northwest Atlantic rocky shores. *Mar Biol Res* 10:51–60
- Väinölä R, Strelkov P (2011) *Mytilus trossulus* in northern Europe. *Mar Biol* 158:817–833
- Wenne R, Zbawicka M, Bach L, Strelkov P and others (2020) Trans-Atlantic distribution and introgression as inferred from single nucleotide polymorphism: mussels *Mytilus* and environmental factors. *Genes* 11:530
- Wiesenthal AA, Timm S, Sokolova IM (2025) Osmotolerance reflected in mitochondrial respiration of *Mytilus* populations from three different habitat salinities. *Mar Environ Res* 205:106968
- Wood SN (2017) Generalized additive models: an introduction with R, 2nd edn. **Chapman and Hall**
- Zbawicka M, Saňko T, Strand J, Wenne R (2014) New SNP markers reveal largely concordant clinal variation across the hybrid zone between *Mytilus* spp. in the Baltic Sea. *Aquat Biol* 21:25–36
- Zolotarev V**, Shurova N (1997) Relations of prismatic and nacreous layers in the shells of the mussel *Mytilus trossulus*. *Russ J Mar Biol* 23:26–31

Editorial responsibility: Rochelle D. Seitz,
Gloucester Point, Virginia, USA
Reviewed by: 3 anonymous referees

Submitted: September 3, 2024
Accepted: August 4, 2025
Proofs received from author(s): • • • • •, 2025