**Results**

**—переписанный абзац см. ниже Geographical variation in the manifestation of mussel morphotypes.** The binary morphological character that we studied was previously defined as the “presence/absence of a distinct uninterrupted dark prismatic strip under the ligament” (Katolikova et al. 2016; Khaitov et al., 2018). Only material from the White Sea was considered in previous studies. While E-morphotypes in all new populations studied looked the same (the srtip absent: the nacreous layer totally or partially covers the space under the ligament nympha, ESM Fig. +), analysis of geographical data revealed some variation among T-morphotypes unseen previously in the White Sea. In rare shells from most geographical populations studied the nacreous-free strip of prismatic layer was quite narrow and looked as a stria rather than a strip (ESM Fig. +). Further, in all the Gulf of Maine T-morphotypes the same as in rare mussels from other populations the color of the strip was pale (as a shell prismatic layer per se) rather than dark. It made T-morphotypes hardly noticeable to the unaided eye. To reveal the T-morphotype unambiguously it is necessary to find using a dissecting microscope the morphologically pronounced scar that defines the boundary of the nacreous layer under ligament nympha (ESM Fig. ++). The most honest diagnoses of T- and E- morphotypes would be the presence/absence of the uninterrupted strip of the prismatic layer under the ligament nympha recognizable by a clear scar separating the strip from the nacreous layer of the rest of shell.

The binary morphological character under consideration was previously defined, based on the material from the White Sea, as the “presence/absence of a distinct uninterrupted dark prismatic strip under the ligament” (Katolikova et al. 2016; Khaitov et al., 2018). In this study, which was based on the material from different geographical zones, E-morphotypes in all the populations looked the same and conformed to our previous description: the strip was absent, and the nacreous layer totally or partially covered the space under the ligament nympha (ESM Fig. +). However, T-morphotypes showed some variation previously unrecorded in the White Sea. Firstly, most populations examined in this study contained, though rarely, shells in which the nacreous-free strip of the prismatic layer was quite narrow and looked like a stria rather than a strip (ESM Fig. +). Secondly, in all T-morphotypes from the Gulf of Maine populations and in rare T-morphotypes from the other populations the strip was not dark but pale, as the prismatic layer itself. In such cases, T-morphotypes were difficult to notice by an unaided eye. They could be unambiguously identified only with the help of a dissecting microscope by the presence of a scar defining the boundary of the nacreous layer under the ligament nympha (ESM Fig. ++).

Therefore, we propose an amended description of the character used to distinguish the E-morphotype and the T-morphotype: the presence/absence of an uninterrupted strip of the prismatic layer under the ligament nympha clearly recognizable by a scar separating the strip from the nacreous layer of the rest of the shell. This description was applicable to all the mussel populations examined in this study. — Возможно, это нужно задвинуть в Результаты.ESM Fig. ++. Mussel morphotype variation. A. E-morphotypes: space under the ligament nympha is totally (left) or partly covered by the nacre (right). B. T-morphotypes: a strip of uncovered prismatic layer under the ligament nympha is dark and wide (typical case for all populations but the Gulf of Maine ones, left) or narrow and recognizable by a scar separating the nacreous layer from the strip of the uncovered prismatic layer only (typical case for American mussels, right).

ESM Fig. ++. Variation in the manifestation of mussel morphotypes. A. E-morphotypes: the space under the ligament nympha is totally (left) or partially (right) covered by the nacre. B. T-morphotypes: a strip of uncovered prismatic layer under the ligament nympha is dark and wide (left; typical of most examined populations) or pale and narrow, recognizable by a scar separating it from the nacreous layer (right; typical of the Gulf of Maine populations).**Associations between morphotypes and species-specific genotypes around the Kola Peninsula**

Patterns of *PT*, *P(T|tros)*, *P(E|edu)*, *P(tros|T)*, *P(edu|E)* variation as functions of *Ptros* among samples representing the White Sea (*WS*), the brackish- (*BL*) and the saline (*BH*) Barents Sea are visualized on Fig. 2 whereas results of regression analysis are summarized in table 1.

Variation patterns of *PT*, *P(T|tros)*, *P(E|edu)*, *P(tros|T)*, *P(edu|E)* as functions of *Ptros* in samples from the White Sea (*WS*), the brackish Barents Sea (*BL*) and the saline Barents Sea (*BH*) are visualized in Fig. 2. The results of the regression analysis are summarized in Table 1.

A significant positive association between the proportions of *M. trossulus* (*Ptros*) and the proportions of T-morphotypes (*PT*) in samples was revealed for all the three sample sets (Model 1, Table +, Fig. +). For *WS* and *BL*, the data points were generally scattered around the Y=X line, while the regression lines approached it closely, indicating a high proportionality between *Ptros* and *PT*. For BH, the data points were scattered above the Y=X line and the regression line lay higher that the regression lines constructed for *WS* and *BL*. This means that in samples with a similar taxonomic structure, the frequencies of T-morphotypes were always higher in the saline localities in the Barents Sea than in the White Sea and the brackish localities in the Barents Sea.

—см. ниже Analysis of variation ? T-morphotype frequencies among subsamples of *M.edulis* (*P(T|tros)*) and *M.trossulus* (*P(T|edu)*) against proportions of *M. trossulus* in samples (*Ptros*) revealed the next patterns (Model 2, Table +, Fig. +). The tendency to higher frequency of T-morphotypes among *M. trossulus* than among *M. edulis* was universal, but in the *WS* and *BL* it was strong (for *Ptros*=0.5, expected differences in morphotype frequencies between species are about 65%) while in the *BH* - quite small (differences 18% for *Ptros*=0.5) due to increased *P(T|edu),* yet also significant (confidential intervals do not overlap for *Ptros*=0.5, Fig. ++).

The analysis of the frequencies of T-morphotypes in subsamples of *M.edulis* (*P(T|tros)*) and *M.trossulus* (*P(T|edu)*) against proportions of *M. trossulus* in samples (*Ptros*) revealed the following patterns (Model 2, Table +, Fig. +). There was a universal tendency towards a higher frequency of T-morphotypes among *M. trossulus* than among *M. edulis*. This tendency was quite strong in *WS* and *BL* (expected differences in morphotype frequencies between species about 65% for *Ptros*=0.5). In *BH* it was rather weak (expected differences 18% for *Ptros*=0.5) due to an increased *P(T|edu)* but significant (confidential intervals for *Ptros*=0.5 did not overlap, Fig. ++).

In all three subsets a positive correlation of *P(T|tros)* and *P(T|edu)* with *Ptros* was found, that is with increasing contribution of *M. trossulus* to samples frequencies of T-morphotypes increased both among *M. edulis* and *M. trossulus*. As a result, T-morphotype frequencies among both genotypes were usually few dozens of percent higher in M*. trossulus*-dominated samples than in *M. edulis*-dominated samples. Тут что-то не то, растолкуйте, плиз!

A positive correlation of *P(T|tros)* and *P(T|edu)* with *Ptros* was found in all the three subsets. This means that with the increasing contribution of *M. trossulus* to the samples the frequencies of T-morphotypes increased both among *M. edulis* and among *M. trossulus*. As a result, \*\*\*.The probability of correct identification of *M. trossulus* by the T-morphotype (the frequency of *M. trossulus* among T-morphotypes, P(tros|T) expectedly increased with the increasing Ptros, while the probability of correct identification of *M. edulis* by the E-morphotype (P(edu|E)) demonstrated an opposite pattern (Model 3, Table +, Fig. +). In the M. trossulus - dominated populations, P(tros|T) tended to one (any mussel with a T-morphotype is 100% M. trossulus), while P(edu|E)) tended to zero (any mussel with an E-morphotype is 100% M. trossulus), and vice versa. In the well-mixed samples (Ptros = 0.5) the predictive values for both species was about 0.75-0.85 in WS and BL but only 0.6 - 0.7 in BH (Fig. ?). It means that the morphotype test has a much lower predictive value in the saline Barents Sea than in the brackish Barents Sea and in the White Sea (the predictive value of 0.5 means a random association between the genotype and the morphotype). It is evident from Fig. 2 that a low predictive value of the test in BH is mainly due to a generally low *P(tros|T)*: even though the great majority of *M. trossulus* have a T-morphotype, it is difficult to recognize them because many *M. edulis* have this morphotype, too. On the other hand, E-morphotypes, let??? which are??? not that common in BH samples, are predominantly found in *M. edulis*. Nevertheless, the statistical analysis indicates that both *P(tros|T)* and *P(edu|E)* predicted by the model were smaller in BH than in WS and BL.

For each of the GLMM models considered (Model 2 and 3), marginal and conditional pseudoR2 were close to each other (Table ++). This indicating the weak role of random factor (sample) as regulator of models, i.e. the satisfactory reproducibility of results in different populations~~.~~ This indicates that the role of the random factor (sample) as regulator of models was weak, i.e. the reproducibility of the results in different populations was satisfactory.

In intra-set comparisons, the regression coefficients did not differ statistically for *WS* and *BL* sets, while *BH* was always different from *WS* (Table 1). To assess the possibility of pooling the data sets, we compared the AIC of full Model 3 (AIC = ) with AICs of three other models based on differently pooled WS, *BL* and *BH* sets. The model based on pooled *WS* and *BL* (*WSBL*) and *BH* showed the lowest AIC. Therefore, in the following analyses we will consider two sets, *WSBL* and *BH*.

Figure 2. Variation of PT, P(T|tros), P(E|edu), P(tros|T), P(edu|E) as functions of Ptros in the White Sea (WS), brackish Barents Sea (BL) and saline Barents Sea (BH). Points - empirical estimates, their size is proportional to sample size. Lines – regression model predictions, grey filling – 95% confidence intervals of regressions. (A) Proportions of T-morphotypes (PT) (Model 1). (B). Proportions of T-morphotypes among M. trossulus (filled points) and M. edulis (empty points) (Model 2). (C) Frequencies of M. trossulus among T-morphotypes (filled points) and of M. edulis among E-morphotypes (Model 4). Vertical lines on B and C connect subsamples of M. trossulus and M. edulis from the same samples.

Table 1. Parameters of the fitted regression models.

**Associations between morphotypes and species-specific genotypes around the Atlantic**

The patterns of *Ptros* variation against *PT* and the patterns of *P(T|tros)*, *P(E|edu)*, *P(tros|T)* and *P(edu|E)* variation against *Ptros* in samples from different geographical zones are visualized in Fig. 3. The results of the regression analysis are summarized in Table 1. The Scottish material was not included in the regression analyses. Re-analyses of the data from the White and the Barents Sea (*WSBL* and *BH* sets) together with the data from other regions revealed the same patterns as those described above. Again, in all the cases when mixed models were used (Model 5, Model 6, Table 1) the marginal and conditional pseudoR2 were close to each other (Table ++) indicating a weak role of the random factor (subset) as regulator of models, i.e. a satisfactory reproducibility of the results from population to population in all the regions.

The proportion of *M. trossulus* in samples (Ptros) was positively correlated with the proportion of T-morphotypes (*PT*) in the other sets, as it did in the samples from the White and the Barents Sea. This tendency was significant for all the sets (Fig. 2; Model 4, Table 1). Otherwise, the patterns of variation were different for different sets. For *GOM*, the regression line stretched above the Y=X line but close to it, indicating the proportionality between *PT* and *Ptros*. For *Balt*, the regression slope was very steep, and the regression line rapidly diverged from the Y=X line. This was due to the fact that the *PT* range in *Balt* was, unlike the situation in the other sets, very narrow (0-0.4) as compared with the *Ptros* range (~0-1), and the small surplus of T-morphotypes in the samples was accompanied by a strong increase in the *M. trossulus* prevalence. A similar tendency was observed in the scanty material from *Norw*. Both *SCOT* samples fell on the Y=X line. Noteworthy are single “outlier” samples from *GOM* and *NORW*, in which *PT* was close to zero but *Ptros* washigh. While *P(T|edu)* estimates were low everywhere, in *BH*, *P(T|tros)* demonstrated a strong variation among sets and a noticeable variation within some sets (Fig. 2; …; Table 1). Similarly to *WSBL*, most *M. trossulus* had T-morphotypes in *GOM* and *Sсot* but not in *BALT* and *NORW*. For *Ptros*=0.5, expected differences in the morphotype frequencies between the species were about 44% for *GOM*, 6% for *Balt* and 24% for *Norw*. A significant positive dependence of the frequencies of T-morphotype on *Ptros* among conspecific genotypes, which was so prominent in the White and the Barents Sea, was recorded elsewhere only in *BALT* for *P(T|tros)* (Table 1).

The structure of the dependence of *P(tros|T)* and *P(edu|E)* on *Ptros* in GOM, BALT and NORW (Model 6. Fig. ++, Table +) was the same as in the samples from the Kola Peninsula (Model 3. Fig. +, Table +): *P(tros|T)* increased with the increasing Ptros, while *P(edu|E)* showed an opposite tendency. To simplify and formalize the comparison, we provide the predictions of Model 6 for equally mixed populations (*Ptros*=0.5) together with their 95% confidence intervals in Table 2, where actual proportions of *M. trossulus* among T-morphotypes (*P(T|tros)*) and *M. edulis* among E-morphotypes (*P(T|edu)*) in pooled samples from the respected sets are also provided.

Table 2. Proportions of *M. trossulus* among T-morphotypes (*P(tros|T)*) and proportions of *M. edulis* among E-morphotypes (*P(edu|E)*) in pooled samples (direct count) and in equally mixed samples (predictions by the regression Model 6) in different sample sets. Low and upper boundaries of 95% confidence intervals are provided for predicted values (in brackets).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | P(edu|E) | | P(tros|T) | |
| Set | Ptros=0.5 | In the data | Ptros=0.5 | In the data |
| WBL | 0.77 (0.73-0.81) | 0.86 | 0.85 (0.82-0.89) | 0.86 |
| BH | 0.70 (0.61-0.78) | 0.84 | 0.57 (0.51-0.63) | 0.48 |
| GOM | 0.66 (0.54-0.77) | 0.86 | 0.86 (0.68-0.95) | 0.80 |
| BALT | 0.51 (0.44-0.58) | 0.46 | 0.82 (0.58-0.94) | 0.93 |
| NORW | 0.64 (0.53-0.74) | 0.51 | 0.86 (0.68-0.95) | 0.93 |
| SCOT | - | 0.90 | - | 0.96 |

For equally mixed populations the predictive values of *P(edu|E)* in *BALT* did not differ statistically significantly from 0.5, which corresponds to an equal probability of correct and incorrect identification. *P(edu|E)* in *BH*, though quite low, was statistically significantly different from 0.5*.* At the same time, the probabilities of correct identification of *M. trossulus* by the T-morphotype in *GOM*, *Balt* and *Norw* were quite high (for the range of Ptros>0.5). In general, high predictive values for both species were revealed in WSBL.

Variation in morphotype frequencies between *M. edulis* and *M. trossulus* within and between contact zones revealed in the study is illustrated in Fig. 1, where the estimates of *P(T|edu)* and *P(T|tros)* in pooled samples from different sets are provided. *P(T|edu)* was 53% in the saline Barents Sea and less than 10% in all the other sets. In its turn, *P(T|tros)* was 17% in Western Baltic, 42% in Western Norway, 49% in the Gulf of Maine and more than 75% in the White and Barents Seas and Northern Scotland. *P(T|tros)* estimates in Norway and the Gulf of Maine were much affected by the outlier samples (see above). If we discard these samples, *P(T|tros)* will make up 54% in Norway and 71% in the Gulf of Maine.

Fig. 1 also shows the morphotype frequencies in putatively pure populations of species out of the contact zones studied. Within the ancestral range of *M. trossulus* in the Pacific, the populations were nearly monomorphic for the T-morphotype. In the Gulf of St. Lawrence *P(T|tros)* was 0.81, i.e. close to that in most of the conspecific — в смысле, M.trossulus ?? populations in the Gulf of Maine. All reference *M. edulis* populations from temperate areas (Long Island Sound and Cape Cod in Western Atlantic, Northern and Norwegian Seas in Europe) were nearly monomorphic for the E-morphotype. At the northeast extreme of the species range in East Atlantic, in the Southwestern Barents Sea, *P(T|edu)* varied considerably between samples, in particular between samples from brackish (range 0-3%) and saline (0.35-0.70%) localities (see ESM Table 2), as it did along the Barents sea coast of the Kola Peninsula. Increased *P(T|edu)* was also recorded in two northernmost samples from Western Atlantic (both from saline localities), Greenland (66%) and the Gulf of St. Lawrence (73%).

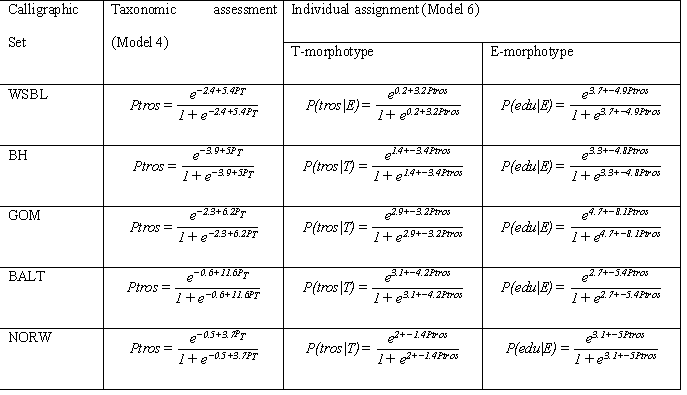
**Associations between morphotypes and shell size**

There was no clear statistical relationship between the size and the morphotype of conspecific mussels. At the level of individual samples, the probability of finding a T-morphotype increased with the mussel size (a positive slope-term of the regression) in 16 out of 34 informative comparisons (when species-specific genotypes were both present and polymorphic for morphotypes) for *M. edulis* and in 17 out of 43 comparisons for *M. trossulus*. The slope-terms of the regression models were individually significant (p<0.05) in four cases for *M. edulus* and in four cases for *M. trossulus*, but only in one case when the correction for multiple testing was applied (sample Bergen\_MV, see ESM1). We also checked the presence of any patterns in residuals from Model 6 as a function of mussel size.

**Prediction of taxonomic structure of populations and predictive values of the morphotype test basing on calibrating samples**

Using the coefficients of the regression models Model 4 and Model 6 (Table 1), we constructed a set of formulas predicting the taxonomic structure (Ptros) and the probability of correct species identification using the morphotype test (Table 3). These formulas were used for the comparison of predictions made with these regression models and the predictions proposed by Eq. 1, 2 and 3.

Table 3. **Formulas used for taxonomic and individual assignment using morphotype tests in different sample sets**



We applied Eq.1 and Eq. 2 (predictive values as a function of *Ptros*, *P(T|tros)* and *P(E|edu)*, “genotype by morphotype calculator”) and Eq. 3 (“*Ptros* by *PT* calculator”) using as an input the data on all possible pairs of populations from *WSBL* and compared the values predicted by these equations with those predicted by regression models 6 and 4, respectively (Table 3, новая таблица с формулами). Fig. 4 illustrates the goodness of correspondence of the tw predictions depending on the genetic constitution of the paired samples as expressed by the Delta index. The best predictions of *Ptros* were obtained when the most dissimilar samples consisting of nearly pure *M. edulis* and *M. trossulus* were used, while the best predictions of P(edu|E) and P(tros|T) values were obtained when the most mixed samples (*Ptros* of both samples close to 0.5) were taken for calibration. We applied the “calculators” to all five geographical sets using, where possible, two most dissimilar samples for the calculation of *Ptros* and two most mixed samples for the calculation of predictive values (Fig. 3; note that only two samples were available for *Scot*). Visual inspection of Fig. 3 shows good correspondence between the predicted by the “genotype by morphotype calculator” and regression lines не поняла, между чем и чем? in all cases but in *NORW*. The latter was due to the formal choice of the only outlier sample with an extremely low *P(tros|T)* as a calibrating one. In its turn, the “*Ptros* by *PT* calculator” was inaccurate for *BH*, *NORW* and *BALT* but nearly ideal for *WSBS* and *GOM*.

Fig. +. Correspondence between “genotype by morphotype calculator” (Eq. 1-2 , left graph) and “*Ptros* by *PT* calculator” (Eq. 3, right graph) and regression models (Model 6 and Model 4, respectively). Each point corresponds to a unique pair combination of samples from WSBL. . OX axis reflects dissimilarity of genetic structure in each pair (Delta) (for pure conspecific samples Delta takes a value of zero, for equally mixed samples – 0.5, for two pure heterospecific samples - 1). OY: goodness of correspondence between assessment of predictive values by equations and regression models.