Gene 876 (2023) 147517

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

The molecular background of the aspartate aminotransferase polymorphism in *Littorina* snails maintained by strong selection on small spatial scales

Felix Mittermayer [a,c], Cecilia Helmerson [b,c], Mårten Duvetorp [c], Kerstin Johannesson [c], Marina Panova [c,*]

[a] Research Division Marine Ecology, Research Unit Marine Evolutionary Ecology, GEOMAR Helmholtz Centre for Ocean Research Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany
[b] Centre for Ecological and Evolutionary Synthesis, Department of Biosciences, University of Oslo, Norway
[c] Department of Marine Sciences, University of Gothenburg, Tjärnö, 45296 Strömstad, Sweden

ARTICLE INFO

Edited by: Isabel Mauricio, PhD

Keywords:
Allozyme
Littorina saxatilis
Marine snail
Gene duplication
Gene structure
Genetic clines
Selection

ABSTRACT

Allozymes present several classical examples of divergent selection, including the variation in the cytosolic aspartate aminotransferase (AAT) in the intertidal snails *Littorina saxatilis*. AAT is a part of the asparate-malate shuttle, in the interidal molluscs involved in the anaerobic respiration during desiccation. Previous allozyme studies reported the sharp gradient in the frequencies of the $AAT^{100}$and the AAT^{120} alleles between the low and high shores in the Northern Europe and the differences in their enzymatic activity, supporting the role of AAT in adaptation to desiccation. However, the populations in the Iberian Peninsula showed the opposite allele cline. Using the mRNA sequencing and the genome pool-seq analyses we characterize DNA sequences of the different AAT alleles, report the amino acid replacements behind the allozyme variation and show that same allozyme alleles in Northern and Southern populations have different protein sequences. Gene phylogeny reveals that the AAT^{100} and the northern AAT^{120} alleles represent the old polymorphism, shared among the closely related species of *Littorina,* while the southern AAT^{120} allele is more recently derived from AAT^{100}. Further, we show that the Aat gene is expressed at constitutive level in different genotypes and conditions, supporting the role of structural variation in regulation of enzyme activity. Finally, we report the location and the structure of the gene in the *L. saxatilis* genome and the presence of two additional non-functional gene copies. Altogether, we provide a missing link between the classical allozyme studies and the genome scans and bring together the results produced over decades of the genetic research.

1. Introduction

Divergent selection favouring alternative alleles in different environments maintains stable polymorphisms in populations (Hedrick et al., 1976; Nosil et al., 2009; Mojica et al., 2012; Comeault et al., 2015; Laurent et al., 2016; Tigano and Friesen, 2016). The classical studies of selection on allozyme variation show many examples of allelic variation strongly associated with contrasting habitats (e.g., Watt, 1977; Koehn et al., 1983; Crawford and Powers, 1989; Hummel et al., 1997). Allozyme polymorphisms are also among relatively few examples when we can connect substitutions at DNA level through changes in protein properties and physiological reactions to the fitness, as in the case of alcohol dehydrogenase polymorphism and adaptations to food substrates in *Drosophila* (McDonald and Kreitman, 1991), phosphoglucose isomerase polymorphism and flight performance at different temperature in *Colias* butterflies (Wheat et al., 2006) and lactate dehydrogenase in temperature adaptation of killifish (Powers and Schulte, 1998).

A first step in generating hypotheses regarding the key molecular mechanisms involved in adaptation is to characterize the sequences of the different alleles (Dean and Thornton, 2007; Dalziel et al., 2009; Storz

Abbreviations: AAT, aspartate aminotransferase, protein; *Aat*, aspartate aminotransferase, DNA/gene; cDNA, complimentary DNA; cds, coding sequence; EPIC primer, exon primed intron crossing primer; IEP, isoelectric point; mya, million years ago.

* Corresponding author.
E-mail address: marina.panova@marine.gu.se (M. Panova).

https://doi.org/10.1016/j.gene.2023.147517
Received 17 March 2023; Received in revised form 14 May 2023; Accepted 24 May 2023
Available online 29 May 2023

and Wheat, 2010; Andersen, 2012). While genome scans provide a whole-genome perspective on the divergence under selection, it is difficult to separate the target of selection and the linked hitchhiking loci within the divergent genome region (Bierne et al., 2011; Ravinet et al., 2017). Further, the resolution of genome scans is often not high enough to provide the information on variation in each nucleotide, and non-synonymous substitutions are easily missed. For example, a recent whole-genome sequencing at 5X coverage of the two ecotypes of the snail *Littorina fabalis* failed to spot several non-synonymous mutations in the arginine kinase, another allozyme locus showing environmental clines in allele frequencies, simply because those sites did not have any coverage after filtering (Le Moan et al., 2022).

In this study we use direct sequencing to characterize allelic variation in another classical allozyme system, aspartate aminotransferase (AAT) showing strong clinal variation between the shore levels in the marine snail *Littorina saxatilis* (Olivi, 1792) (Johannesson et al., 1995). This species has become a model system to study ecological divergence (Johannesson et al., 2017; Ravinet, 2018), facilitated by the recently sequenced genome for the species (Westram et al., 2018). Several recent genomic studies have been focusing on the Crab and Wave ecotypes of *L. saxatilis* and revealed divergence in multiple regions across the genome, with approx. 50% of the variation coupled to chromosomal inversions (Westram et al., 2018; Faria et al., 2019; Morales et al., 2019; Koch et al., 2021, 2022). Interestingly, one earlier study that looked at the genome-wide variation over the vertical shore gradient found divergence also between low and high shore snails in a part of the genome (Morales et al., 2019). The vertical environmental gradient is here shaped by the regular tidal variation that exposes *L. saxatilis* snails inhabiting different shore levels to different degrees of stress from desiccation and extreme temperatures. From earlier studies (Johannesson et al. 1995) the aspartate aminotransferase gene seems a strong candidate for divergent selection, and a molecular characterisation of the gene is a first useful step to address its role in adaptation of the snails to the heterogeneous shore environment.

1.1. Nomenclature

Aspartate aminotransferase (AAT, EC 2.6.1.1) is sometimes also denoted aspartate transaminase, AST, or glutamic oxaloacetic transaminase, GOT. It catalyses the reversible conversion of aspartate and α-ketoglutarate to oxaloacetate and glutamate. Eukaryotes possess the cytosolic and the mitochondrial AAT, the latter is encoded by a nuclear gene but acts in mitochondria (Winefield et al., 1995). This study focuses on the cytosolic form; unless it is stated otherwise, hereafter we refer to the cytosolic aspartate aminotransferase. To distinguish the protein and its gene we will use "AAT" and "*Aat*" designations.

1.2. AAT allozyme variation in Littorina saxatilis

Protein electrophoresis the intertidal snail *Littorina saxatilis* has earlier revealed the presence of two alleles of the cytosolic AAT AAT^{120} and AAT^{100} (the numbers refer to the distances they migrate in the gel under standard conditions). The same allozyme alleles were also found in two closely related species *L. arcana* and *L. compressa* (Sundberg et al., 1990; Knight and Ward, 1991). The mitochondrial AAT did not show much allelic variation in *L.* saxatilis and during the applied electrophoresis conditions migrated in the opposite direction (Janson and Ward, 1984).

The frequencies of the AAT^{100} and AAT^{120} alleles in *L. saxatilis* show remarkably steep vertical gradients across the shore (Johannesson and Johannesson, 1989; Johannesson et al., 1993; Johannesson and Tatarenkov, 1997). In northeast Atlantic, the AAT^{100} decreases in frequency from low to high shore level while, somewhat unexpectedly, the opposite trend is observed in Spain (Fig. 1). The allele frequency changes occur at very small spatial scales, ≤10 m (Johannesson and Johannesson, 1989). The strong correlation between the shore level and the AAT allele frequencies as well as a rapid return of the original cline after natural perturbation together provide convincing support for differential selection maintaining the polymorphism in this trait (Johannesson et al., 1995).

Solid lines represent sites from Northeast Atlantic:Norway, two sites from Iceland, UK and Sweden, site names are indicated on the right side (data from Johannesson and Johannesson, 1989). Dashed lines represent three sites from Spain, site names are indicated on the left side (data from Johannesson et al., 1993).

1.3. Aspartate aminotransferase function

Aspartate aminotransferase catalyses the reversible transfer of an amino acid group between aspartate and glutamate and is hence important in the amino acid metabolism and the aspartate-malate shuttle. Aspartate is also used as a substrate in anaerobic metabolism in molluscs (Hochachka, 1980; de Zwaan, 1991). Potentially it is an important energy source for intertidal molluscs exposed to desiccation, and aspartate depletion was observed in *L. saxatilis* during anaerobiosis (Sokolova and Pörtner, 2001).

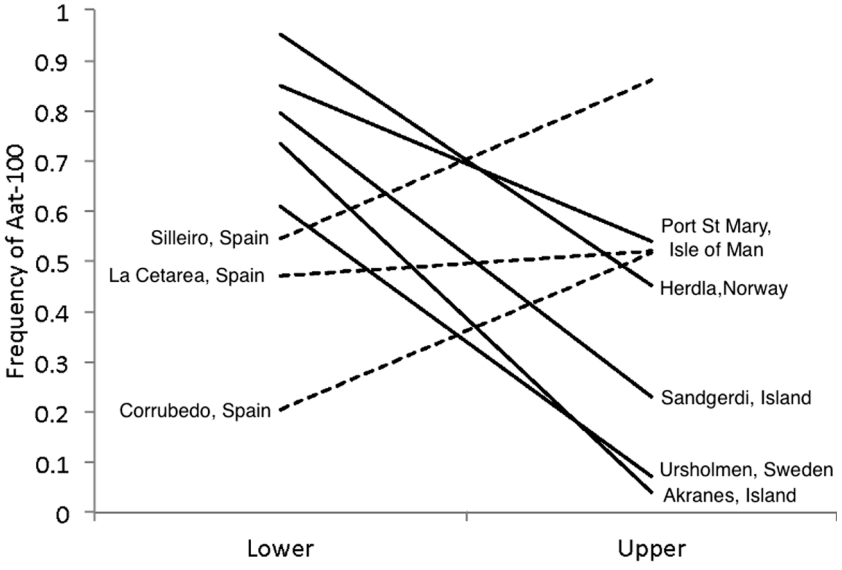

Fig. 1. Clines of the AAT^{100} allele from low to high shore in *L. saxatilis*.

Interestingly, intertidal snails living at the different shore levels demonstrate different metabolic strategies when exposed to desiccation: snails at the low shore maintain the metabolism by having high anaerobic rate while snails at the high shore have low rate of anaerobiosis and metabolism in general (McMahon, 1990; Sokolova and Pörtner, 2001). One plausible explanation is that maintaining normal metabolic rates during relatively short periods of desiccation is beneficial at the low shore while high-shore snails must lower their metabolism to survive prolonged desiccation. In agreement with this hypothesis enzymatic AAT activity, measured in crude tissue homogenates was different between the AAT genotypes of *L. saxatilis* (Panova and Johannesson, 2004). Low shore homozygotes ($AAT^{100/100}$) show significantly higher activity than upper shore homozygotes ($AAT^{120/120}$), and heterozygotes ($AAT^{100/120}$) had intermediate activity.

Overall, while the allozyme data indicate that the aspartate aminotransferase in *L. saxatilis* is under strong selection, the underlying genomic variation is unknown. Further, it is not known whether the difference in enzymatic activity depends on the structural variation of the allozymes or their expression level. In this study, we address these questions with the following main aims:

1. Characterize coding sequences of the different allozyme alleles in *L. saxatilis* from northern and southern European populations. To reconstruct the genealogy of the different *Aat*-alleles we have also included four other North-Atlantic *Littorina* species (Fig. 2): the two closely related species, *L, compressa* and *L. arcana* (separated <2 mya from *L. saxatilis*), the more distantly related *L. fabalis* and *L. obutsata* (separated 2–4 mya), and the distant congeneric *L. littorea* (separated ~15 mya; Reid, 1996; Reid et al., 2012; Panova et al., 2014).
2. Compare the gene expression level in the different AAT genotypes of *L. saxatilis* in order to detect whether previously shown differences in enzymatic activity depend on the expression level or are rather linked to structural protein variation.
3. Based on the coding sequences locate the genes and describe the structure of the cytosolic and mitochondrial *Aat* genes in the *L. saxatilis* genome.

2. Material and methods

2.1. Characterization of the coding sequences for the AAT alleles in Littorina

2.1.1. Amplification and sequencing of the alleles

Samples of *L. saxatilis*, were collected at sites in Norway (Tromsö, Sommaroya, 69° 38′ 43.27″ N, 18° 00′ 04.41″ E), Iceland (Keflavik, Sandgerdi, 64° 01′ 54.97″ N, 22° 42′ 51.22″ W), Sweden (Tjärnö marine laboratory, Yttre Vattenholmen, 58° 52′ 30.64′’ N, 11° 06′ 21.21″ E) and Portugal (Vila Praia de Âncora, 41° 49′ 14.84″ N, 08° 52′ 26.64″ E). At each site the snails were collected at the high and low range of the species distribution (separated by ≤10 m) to ensure that samples included homozygotes for both alleles. Additional samples of other species were collected in Norway, Trömsö (*L. arcana*, and *L. compressa*) and Sweden, Tjärnö Marine laboratory (*L. fabalis* and *L. littorea*) in their typical habitats. The snails were kept alive or stored at −80° C until dissection. Head or foot tissue was stored in RNA*later* for later extraction of RNA. The remaining tissue was homogenized and used for allozyme genotyping by native acrylamid gel protein electrophoresis, adapting the original starch electrophoresis protocol from Tatarenkov and Johannesson (1994). Homozygotes were selected for RNA extraction, followed by singled-stranded cDNA synthesis with Promega reverse transcriptase, resulting in a total of 22 samples representing different allozyme genotypes, localities and species (Table 1).

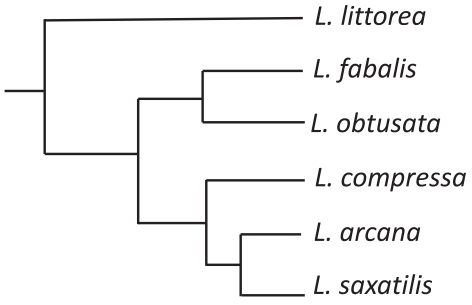

Fig. 2. Dendrogram representing the phylogeny of the six North-Atlantic species of *Littorina*.

Multiple primers were designed to amplify the *Aat* coding sequence (cds) using genome contigs from *L. saxatilis* genome assembly v. 1 (Westram et al. 2018; Dryad repository https://datadryad.org/stash/dataset/https://doi.org/10.5061/dryad.bp25b65) that showed high similarity to the *Aat* protein sequences found in the genome of the Pacific oyster (Boutet et al. 2005), and two partial *L. saxatilis* transcripts from the transcriptome assembly provided by A. Sa Pinto at CIBIO, Portugal. These sequences were aligned with the multiple alignment tool and primers were designed in Geneious 6.1.4. The success of different primer combinations was evaluated by visualising PCR products on the agarose gel electrophoresis. All resulting products of good quality and correct length were purified using ExoSAP-IT or E.Z.N.A Cycle-Pure kit and sent for sequencing to Macrogen Europe (Amsterdam). Resulting sequences were aligned with the cds predicted from the genome and transcriptome sequences.

Among many different primer combinations, the best results were obtained with *Aat*-F2 (5′-ACAGGGATATTCATCGGCAAC-3′) and *Aat*-R9 (5′-CCTGGCACCTTCTTCTGTC-3′) as reverse primers. These primers amplify 1046 bp fragment of the *Aat* cds, corresponding to a large part of the total predicted length of the coding sequence 1230 bp, or 348 of the predicted 410 amino acid residuals, including the start codon. This fragment was amplified using 5 PRIME HotMaster *Taq* Polymerase and a touchdown cycling program (calculated for $T_a = 55°$). The PCR products were individually cloned into *E. coli* with a TOPO TA Cloning kit, colonies were picked and screened using M13 PCR protocol and gel electrophoresis, and 4–12 positive clones per individual were sequenced at Macrogen Europe.

The GeneRacer kit (Invitrogen) was used to amplify and sequence the UTRs of the transcript and the missing parts of the cds, combining the *Aat* gene specific primers above and the GeneRacer primers targeting 5′ capped end and 3′ poly-A tail of the transcripts. Sequences of 3′UTR were used to obtain the C-terminus of the protein, not recovered by the gene-specific primers, and sequences of both 5′ and 3′ UTR were used to locate the functional gene copy in the genome, see below. The PCR products were screened on agarose gel electrophoresis for the fragment of the correct length; when reactions contained multiple fragments, it

Table 1
List of *Littorina* individuals used to characterize the *Aat* coding sequences. The last column shows the total number of sequenced individuals; number in brackets indicate if only partial coding sequences were obtained for some of them.

Species	Country	Shore level	Expected allele	N individuals
L. saxatilis	Sweden	High	AAT^{120}	2
L. saxatilis	Sweden	Low	AAT^{100}	2
L. saxatilis	Iceland	High	AAT^{120}	2
L. saxatilis	Iceland	Low	AAT^{100}	1 (1)
L. saxatilis	Norway	High	AAT^{120}	2 (1)
L. saxatilis	Norway	Low	AAT^{100}	2 (2)
L. saxatilis	Portugal	High	AAT^{100}	2 (1)
L. saxatilis	Portugal	Low	AAT^{120}	2
L. arcana	Norway	High	AAT^{120}	2
L. arcana	Norway	Low	AAT^{100}	1
L. compressa	Norway	Low	AAT^{100}	1
L. fabalis	Sweden	Low	Other	1
L. littorea	Sweden	Low	Other	2

was cut from the gel. These amplified fragments were purified and used as templates for nested PCR. Products from this nested PCR were individually cloned into *E. coli* using a Zero Blunt TOPO PCR cloning kit. Colonies were picked, amplified using M13 PCR protocols and screened on gel electrophoreses. All positive samples containing the product were purified using the Cycle-Pure kit and sequenced at Macrogen Europe.

2.1.2. Sequence data analysis

Forward and reverse sequences produced with the gene-specific primers and the GeneRacer were inspected for quality and assembled using Geneious. Phylogenetic analyses were performed using a Bayesian approach with the program MrBayes (Huelsenbeck and Ronquist, 2001) and haplotype networks were created in TCS 1.21 (Clement et al., 2002). The predicted AAT protein sequence was used to retrieve protein sequences for other molluscs from the NCBI database using the blastp algorithm. These sequences were aligned using Clustal algorithm and a Neighbour-Joining tree based on Jukes-Cantor distances was constructed in Geneious.

The isoelectric point (IEP) describes the pH at which a molecule has no electric charge. Differences in IEP affect the speed at which different enzymes move in a gel with a certain pH during electrophoresis. We estimated the isoelectric points for the different allozyme haplotypes from their predicted protein sequences using Geneious and compared with the migration patterns observed in the gels.

Finally, the obtained coding and UTR sequences were used to search the available *L. saxatilis* transcriptome assemblies with the blastn algorithm: an unpublished assembly from 454 sequencing kindly provided by A. Sa Pinto and published Illumina sequencing assemblies (NCBI TSA: GHPE00000000, GHUM00000000 and GHUL00000000).

2.1.3. Allele frequencies in the pool-seq data

The cloning of individual alleles is laborious and can only be performed on a limited number of individuals. Once we had the SNPs variant distinguishing the alleles, we used the pool-seq data from Morales et al. (2019) to compare the DNA allele frequencies with the previous allozyme data. Morales et al. (2019) sequenced pooled DNA samples from 24 to 48 individuals of *L. saxatilis* from the low and high shore at different sites in Europe: Sweden, UK, France and Spain, in all sites including the two main ecotypes of the species - the crab and the wave ecotype; see Suppl. information in Morales et al. (2019) for detailed description of the sites. At several sites the Crab and Wave ecotypes are vertically distributed: in UK and France the Crab-ecotype is on the low shore and the Wave-ecotype is on the high shore, while in Spain it is the other way around. In Sweden the Crab ecotypes live on boulder beaches while the Wave ecotype live on rocky outcrops; for Sweden we only used data for the Wave ecotype. We also included pool samples of *L. compressa* and *L. arcana* from UK that were sequenced along with *L. saxatilis* but not included in Morales et al. study (NCBI: PRJNA494650). We extracted the nucleotide frequencies for the allele-characteristic sites from the bam-files and based on them calculated the *Aat* allele frequencies in different populations and species of *Littorina*.

2.2. Gene expression in the emerged vs submerged conditions and in different genotypes

Adult snails (4–5.5 mm size) were sampled in January 2014 in the hybrid-zone of the *Aat* cline at an exposed rocky shore on the island Ursholmen at the Swedish west coast (58° 49′ 52.74′’ N10° 59′ 28.20″ E). Snails were initially kept in aquarium with running seawater of ambient temperature and salinity for 84 days. Thereafter one group were exposed to emersion placed on filter paper at room temperature for 48 h while the other group remained in the aquarium. From the group of desiccated snails samples were taken at 0 h, 2 h, 4 h, 6 h, 12 h, 24 h and 48 h; control snails from the aquarium were sampled at 0 h and 48 h. The snails were immediately killed by crushing and preserving in RNAlater solution (Ambion) at −20 °C until RNA extractions.

The samples were thawed on ice, shell and operculum were removed and RNA from the whole snails was extracted with TRI-Reagent (Sigma-Aldrich) following the manufacturer's protocol. The RNA quality and quantity was assessed using NanoDrop ND-1000 Spectrophotometer and denaturing agarose gel electrophoresis, and the RNA was stored at −80 °C. RNA was converted to single-stranded complementary DNA (cDNA) by reverse transcription with iScript Select cDNA synthesis kit (Bio-Rad) following the protocol with random primers.

Gene expression analysis was done using 3–4 snails per treatment and controls (33 snails in total). Quantification of gene expression was performed as comparative C_T in StepOnePlus Real-Time PCR System (Applied Biosystem). Gene expression was measured for the cytosolic and the mitochondrial *Aat* genes and five control genes: 18S, ubiquitin, ribosomal protein 1, histone 3.3 and elongation factor 2. Primers for the *Aat* genes, 18S, ubiquitin and ribosomal protein 1 were designed using Primer Express v. 2.0 (Applied Biosystems) based on sequences of the cytosolic *Aat* from this study and from the *Littorina* sequence database (Canbäck et al. 2012) for other genes, see Suppl. Table S1. The primers for the histone 3.3 and elongation factor 2 were obtained from Martínez-Fernández et al. (2010).

Quantitative PCR reactions consisted of 10 μl Fast SYBR Green Master Mix (Applied Biosystems), 4 μl of cDNA diluted 1:10, 2 μl nuclease-free water and 2 μl of forward primer and reverse primers at 0.5 μM concentrations. Cycling conditions were 95 °C for 20 s, 40 cycles at 95 °C for 30 s and 60 °C for 30 s, followed by a melt curve stage with 95 °C for 15 s, 60 °C for 1 min and 95 °C for 15 s. Each sample was analysed in two replicates and each primer combination had two negative controls.

C_t values were calculated as the mean of the two technical replicates for each individual and gene and corrected by the efficiency values obtained from validation. ΔC_t were calculated in relation to the geometrical mean of the C_t for the control genes. Finally, gene expression levels between the treatments were compared using Kruskal-Wallis ANOVA nonparametric test.

The *Aat* genotypes were determined for the snails after the gene expression analyses. A 477 bp fragment of the *Aat* coding sequence, including 2 sites distinguishing Aat^{100} and Aat^{120} alleles, was amplified using ss-cDNA from the gene expression analysis and the primers *Aat*_genF (5′-TTCAGTAACGTGAAGATGGG-3′) and *Aat*_genR (5′-GTCAACACCCAGATTTTTGG-3′). PCR was performed in 50 μl reaction volume, containing 5 μl 5 PRIME™ HotMaster buffer, 1 μl of dNTPs 10 mM, 1 μl of the forward and reverse primer 10 μM, 0.4 μl of Hot-5 PRIME™ HotMaster™ *Taq* polymerase, 4 μl of the scDNA template and 37.6 μl water. PCR cycling consisted of 4 min initial denaturation at 94 °C, 30 cycles of 94 °C for 30 s, 50 °C for 30 s and 65 °C for 1 min, and final elongation 65 °C for 7 min. PCR products were cleaned with E.Z.N.A Cycle-Pure kit and sequenced by Macrogen. Chromatograms from the forward and reversed sequencing for each sample were assembled and the allele-characteristic sites were carefully inspected to determine genotype. Heterozygotes were called when there were clear double peaks at these sites. To confirm that direct sequencing reliably distinguishes homo- and heterozygotes, a subset of the samples was also genotyped by cloning and sequencing using the cloning protocol above.

2.3. Genome analyses

The predicted coding sequence for the *Aat* gene was used as a query to search for the gene in the *Littorina saxatilis* genome assembly v.1 (Westram et al., 2018). Searches were done using blat (Kent, 2002) and blastn (Altschul et al., 1990) algorithms and exonerate (Slater and Birney, 2005) was used to predict exon–intron boundaries. The 5′ and 3′ UTR sequences were also aligned to the genome using blastn.

3. Results

3.1. DNA sequences of the Aat alleles in Littorina

Full or partial coding sequences for the *Aat* gene were obtained for 22 individuals (Table 1). By assembling the sequences obtained by amplification with the gene-specific primers and the cloning of 3′ UTR we were able to characterize the full cds (1233 bp corresponding to 410 amino acid residuals plus STOP-codon) for nine *L. saxatilis*, three *L. arcana*, one *L. compressa*, one L. *fabalis* and two *L. littorea;* in total 16 individuals. Of them three individuals had two different haplotypes different by synonymous substitutions. By using the gene-specific primers only, we obtained partial coding sequences of 1046 nt (missing 62 residuals at C-terminus) for additional six *L. saxatilis* individuals, representing different allozyme genotype (Table 1). All cds variants are deposited in the NCBI: OQ863267 – OQ863291.

Sequencing of the cloned 3′ UTRs using the GeneRacer technique returned two very different sequence variants per individual; in the following genome analyses they were identified to belong to two different copies of the gene, one functional and one non-functional, see below. For the cds assemblies we chose the variants corresponding to the functional gene copy (the cds obtained by sequencing also correspond to this copy). For this copy, two *Aat* haplotypes (different by synonymous substitutions) were extracted from the genome assembly and included into the analyses. Based on the allele-characteristic sites the sequenced individual (the Crab-ecotypes from the low shore) was a homozygote for AAT^{100}.

Comparison of translated sequences corresponding to the different allozyme alleles showed that all AAT^{100} had the same protein sequence but AAT^{120} alleles were divided into two groups: one including all northern (Scandinavian) samples and southern (Portugal/Iberian Peninsula) samples. Both AAT^{120} variants differed from the AAT^{100} by two replacements, but these were present at different positions: Q/E in residual 29 plus I/V in residual 102 in the northern variant and T/A in residual 208 plus Q/K in residual 210 in the southern variant. Consequently, the two AAT^{120} proteins were different from each other by four amino acid residuals (Table 2). Despite these differences both AAT^{120} alleles had the same isoelectric focusing point at pH 8.11 and therefore are not distinguishable in the allozyme electrophoresis (Table 2). From now onward we denote the two Aat^{120} variants as northern (Aat^{120}-N) and southern (Aat^{120}-S).

Aat^{120}-N and Aat^{120}-S are Aat^{120} alleles from the northern (Scandinavia) and southern (Portugal) populations. Replacements distinguishing each of the two Aat^{120} alleles from Aat^{100} are highlighted. The mitochondrial *Aat* is included for comparison of the isoelectric points (IEP). At the applied electrophoretic conditions (pH = 8.6) all cytosolic AAT alleles are expected to migrate to the minus and the mitochondrial AAT to the plus pole.

Littorina arcana and *L. compressa* shared the allozyme alleles with the *L. saxatilis*. Three analysed individuals of *L. arcana* had both the Aat^{100} and Aat^{120} -N variants. The only individual of *L. compressa* that we sequenced was homozygote for Aat^{100}. Considering synonymous substitutions, we found five Aat^{100} haplotypes, three Aat^{120}-N and two Aat^{120}-S haplotypes in the three species (Fig. 3).

The sequenced haplotypes of the more distantly related species *L. fabalis* and *L. littorea* differed with 8 and 44 nucleotides, respectively from one of the Aat^{100} haplotypes found in *L. saxatilis*. Of these, 3 and 14, respectively were non-synonymous mutations. The phylogenetic reconstruction using the Bayesian approach and *L. littorea* as outgroup (Fig. 4) returned two clades: one containing Aat^{120}-N haplotypes and another containing Aat^{100}, Aat^{120}-S and *L. fabalis* haplotypes. This suggests that the two clades represent an old polymorphism, and one was retained in *L. fabalis*. However, only one individual of *L. fabalis* was sequenced since this species did not show the allozyme AAT polymorphism in the previous studies (Tatarenkov and Johannesson, 1994). The Aat^{120}-S appears to be a more recently derived variant evolved from the Aat^{100} allele.

Different *Aat* alleles are colour-coded: Aat^{100} (blue), Aat^{120}-N (brown), Aat^{120}-S (yellow). The numbers are posterior probabilities for the clades.

The analysis of the pool-seq data confirmed the expected allele frequencies in different populations of *L. saxatilis* and its close relatives. In Sweden, France and UK the low shore-samples of *L. saxatilis* had predominantly the Aat^{100} allele and the Aat^{120}-N appeared in the high shore samples (Table 3). The low shore-species *L. compressa* had only Aat^{100} while *L. arcana* had about equal frequencies of both Aat^{100} and Aat^{120}-N alleles (Table 3). The Aat^{120}-S was only present in the samples from Spain, although it did not show the expected frequency gradient between the low and high shore (Table 3). Interestingly, the high-shore sample from the northern site in Spain (Burela) had both Aat^{120}-N and Aat^{120}-S alleles.

The allele frequencies in the pairs of the sites, characteristic for the different *Aat* alleles were fairy similar, especially for the second pair, where the sites are close and were often covered by the same read. An exception is the two sites distinguishing the alleles in the high shore *L. saxatilis* in Sweden that showed different allele frequency, although both sites indicated the presence of the high-shore allele. This could be due to a sampling effect in the pool-seq data – this sample had the lowest sequence coverage, lower than the number of gene copies in the pool (48) and the estimated allele frequencies may not be accurate.

3.2. Gene expression of the Aat in the emerged vs submerged conditions and in different genotypes

There were no significant differences between the gene expression levels in different treatments neither for the cytosolic ($p = 0.184$) nor for the mitochondrial ($p = 0.708$) *Aat* genes (Suppl. Fig. S1 a,b). The genotyping showed that 12 snails were homozygous for Aat^{100}, 5 snails were homozygous for Aat^{120} and 15 were heterozygotes. Since there was no significant effect of the treatment, the gene expression level of the cytosolic *Aat* in the different genotypes was compared across the treatments using a single-factor ANOVA. In the result there was no significant differences between the gene expression levels in the different genotypes ($p = 0.540$, Fig. 5).

3.3. Gene structure of Aat in the L. saxatilis genome and comparison to other molluscs

The blast and blat searches in the *L. saxatilis* genome v.1 using the full coding sequence of the cytosolic *Aat* as a query returned one contig (contig 7214, approx. 74 Kbp long) containing two copies of the *Aat* gene (Fig. 6). Both gene copies consisted of 9 exons that translated into a full protein sequence with start and stop codons and without any frameshift or premature stop mutations. The exons in the first gene copy had 99–100% nucleotide similarity to the characterised coding sequence of the Aat^{100} allele. All introns started and ended with GT/AG splicing signals.

The second copy had only on average 90% nucleotide sequence similarity to the first copy and the coding sequences obtained by cloning. The predicted protein sequences also had only 89% identity, with especially high divergence in the first exon with 9/39 amino acid residuals being different between the copies and an insertion of three amino acid residuals in the second copy. Blastn search of the

Table 2
Non-synonymous sites, molecular weights and isoelectric point of the *Aat* alleles.

Allele	Res. 29	Res. 102	Res. 208	Res. 210	MW, kDa	IEP
AAT^{120} N	E = Glu	V = Val	A = Ala	K = Lys	45.789	8.11
AAT^{120} S	Q = Gln	I = Ile	T = Thr	Q = Gln	45.849	8.11
AAT^{100}	Q = Gln	I = Ile	A = Ala	K = Lys	45.803	8.44
mt-AAT					47.034	9.16

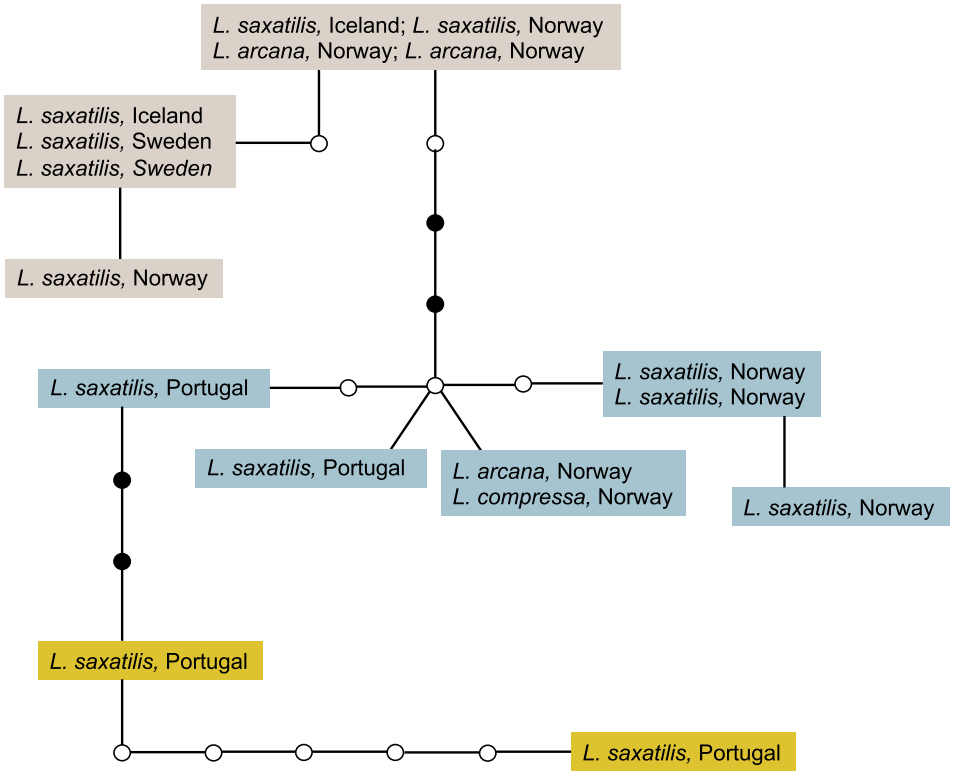

Fig. 3. Network for *Aat* haplotypes found in *L. saxatilis*, *L. arcana* and *L. compressa*. Different *Aat* alleles are colour-coded: Aat^{100} (blue), Aat^{120}-N (brown), Aat^{120}-S (yellow). Synonymous substitutions are denoted by empty circles and nonsynonymous substitutions by filled circles. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

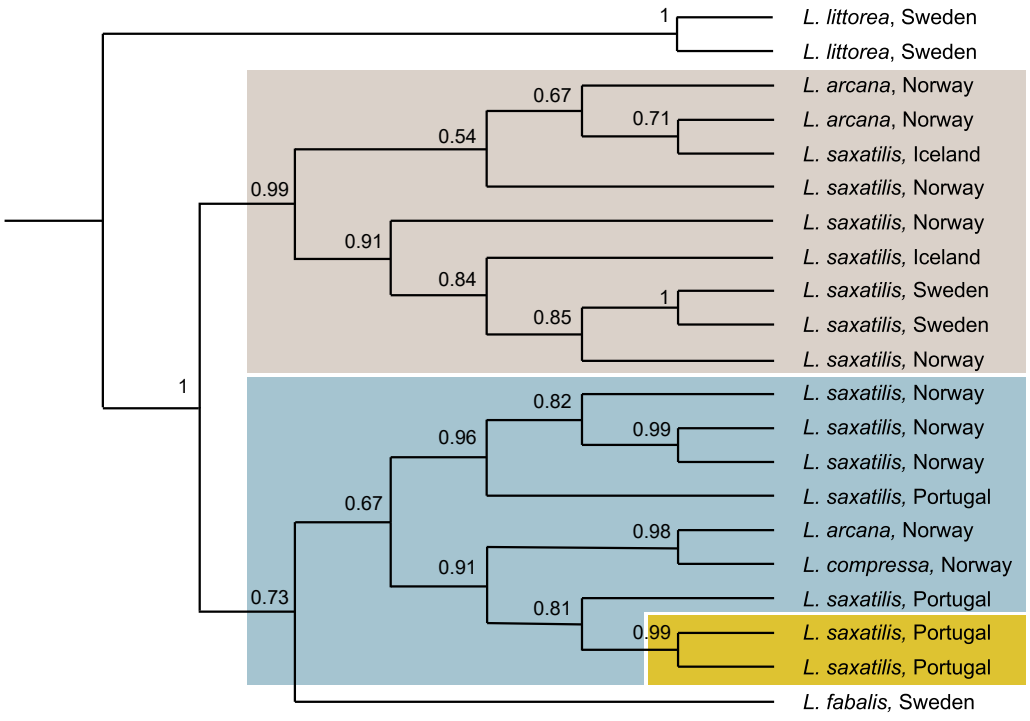

Fig. 4. The Bayesian gene tree for the cytosolic aspartate aminotransferase based on nucleotide sequences in five species of *Littorina*.

transcriptome assemblies returned several transcripts with high similarity to the first gene copy but none similar to the second copy. The 5′UTR and one of the 3′UTR variants obtained from the transcripts by the GeneRacer were found in the genome flanking the coding sequence of the first gene copy but not the second. From this we concluded that the second gene copy is silent (not expressed).

Several introns from the two copies differed in length (Fig. 6), and the introns that had similar length had only 45% similarity between the

Table 3
Aat allele frequencies in different populations of L. saxatilis and its sister-species estimated from pool-seq data from Morales et al. (2019). Nucleotide frequencies in the four sites characteristic for the allozyme alleles in the Table 2 are extracted from the bam-files from the whole genome sequencing of the snail pools (24–48 individuals per pool). For each site, a position in the genome contig 7214, containing the Aat gene, and alternative nucleotide states, corresponding to the different allozyme alleles, are provided in the table. The shaded area highlights the shift from the 120-Northern allele to the 120-Southern allele.

	Position 39,778 G/C 100/120-N	Position 35,361 T/C 100/120-N	Position 25,046 C/T 100/120-S	Position 25,040 T/G 100 /120-S	Average allele frequencies
L saxatilis Low shore Sweden	71/8	71/1	48/0	48/0	100: 0.942 120-N: 0.058 120-S: 0
L saxatilis High shore Sweden	9/17	17/9	29/0	30/0	100: 0.500 120-N: 0.500 120-S: 0
L. saxatilis Low shore France	88/2	113/0	104/0	107/0	100: 0.989 120-N: 0.011 120-S: 0
L. saxatilis High shore France	26/80	36/59	105/0	108/0	100: 0.312 120-N: 0.688 120-S: 0
L. saxatilis Low shore UK	93/17	102/7	115/0	111/0	100: 0.891 120-N: 0.109 120-S: 0
L. saxatilis High shore UK	77/36	81/26	106/0	98/0	100: 0.719 120-N: 0.281 120-S: 0
L. compressa Low shore UK	80/0	81/0	81/0	78/0	100: 1.000 120-N: 0 120S: 0
L. arcana High shore UK	64/70	44/47	126/0	123/0	100: 0.481 120-N: 0.519 120-S: 0
L. saxatilis High shore Spain (Burela)	96/5	49/32	96/23	97/22	100: 0.631 120-N: 0.180 120-S: 0.189
L. saxatilis Low shore Spain (Burela)	77/0	97/0	69/9	76/9	100: 0.889 100-N: 0 120-S: 0.110
L. saxatilis High shore Spain (Silleiro)	92/0	84/0	39/57	40/50	100: 0.425 120-N: 0 120-S: 0.575

two copies. To confirm the presence of these two divergent gene copies in the genome (which could be otherwise an artefact of the genome assembly) we designed two pairs of exon primed intron crossing (EPIC) primers that span across intron 4 (approx. 300 bp) and intron 5 (approx. 500 bp) in the two gene copies. DNA of the snail used in the genome project was amplified with the EPIC primers and the PCR products were cloned and sequenced. This produced two sequence variants that mapped perfectly to the two *Aat* gene copies in the genome and thus confirmed the presence of the two diverged gene copies. Examining the characteristic sites for the different alleles (Table 2) showed that the second copy had P instead of E or Q in the first site; the rest of the sites corresponded to the Aat^{100} and the northern Aat^{120} alleles. In the protein tree the second gene copy was placed as an outgroup to functional (obtain from mRNA) *Aat* copies in *L. saxatilis*, *L. fabalis* and *L. littorea* (Suppl. Fig. S2).

In addition, another truncated copy of the gene was found in the contig 46643. It contains exons 7–9 and another variant of 3′ UTR sequence found in the GeneRacer analyses as well as in the transcriptome assemblies. No other exons were present in the remaining 95 Kbp upstream of the contig, and the contig contains another gene on the opposite strand (Suppl. Fig. S3).

The predicted protein sequence for the cyt-*Aat* had the highest similarity to the cyt-*Aat* in the gastropods *Batillaria attramentaria* (accession number KAG5698065, 77%), *Pomacea canaliculata* (accession number PVD30382, 70% similarity) and *Haliotis rufescens* (accession number XD_048249253, 68% similarity), followed by other molluscs. The length of the cytosolic AAT protein in the molluscs vary between 402 and 416 amino acids, and the predicted protein in *L. saxatilis* of 410 amino acid is within the range. The blastp search also identified the AAT-characteristic protein domains.

Finally, the coding sequence of the mitochondrial *Aat* gene was obtained from transcript TR122003_c0_g1-i2 from the L. saxatilis transcriptome assembly (NCBI: GHUL00000000). Translated protein sequence of this transcript showed 73–78% similarity to the mitochondrial AAT in molluscs, with highest similarity (78,84%) to the protein in the gastropod *Pomacea canaliculata* and the same length (430 aa) as in this species. The identity between mitochondrial and cytosolic AAT in L. saxatilis was 42,08%, and the two forms clustered with cytosolic and in mitochondrial proteins in other molluscs (Suppl. Fig. S2). The mitochondrial *Aat* gene was found on the genome contig 5334.

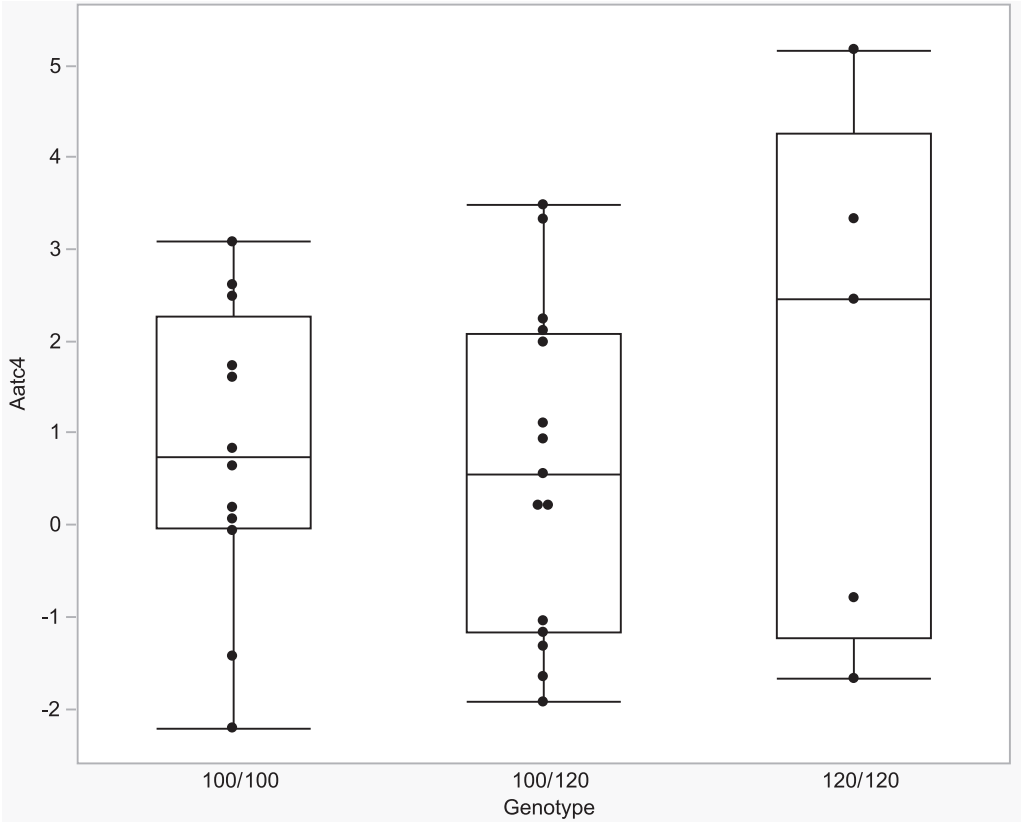

Fig. 5. Expression level of the cytosolic *Aat* in the different genotypes of *Littorina saxatilis*.

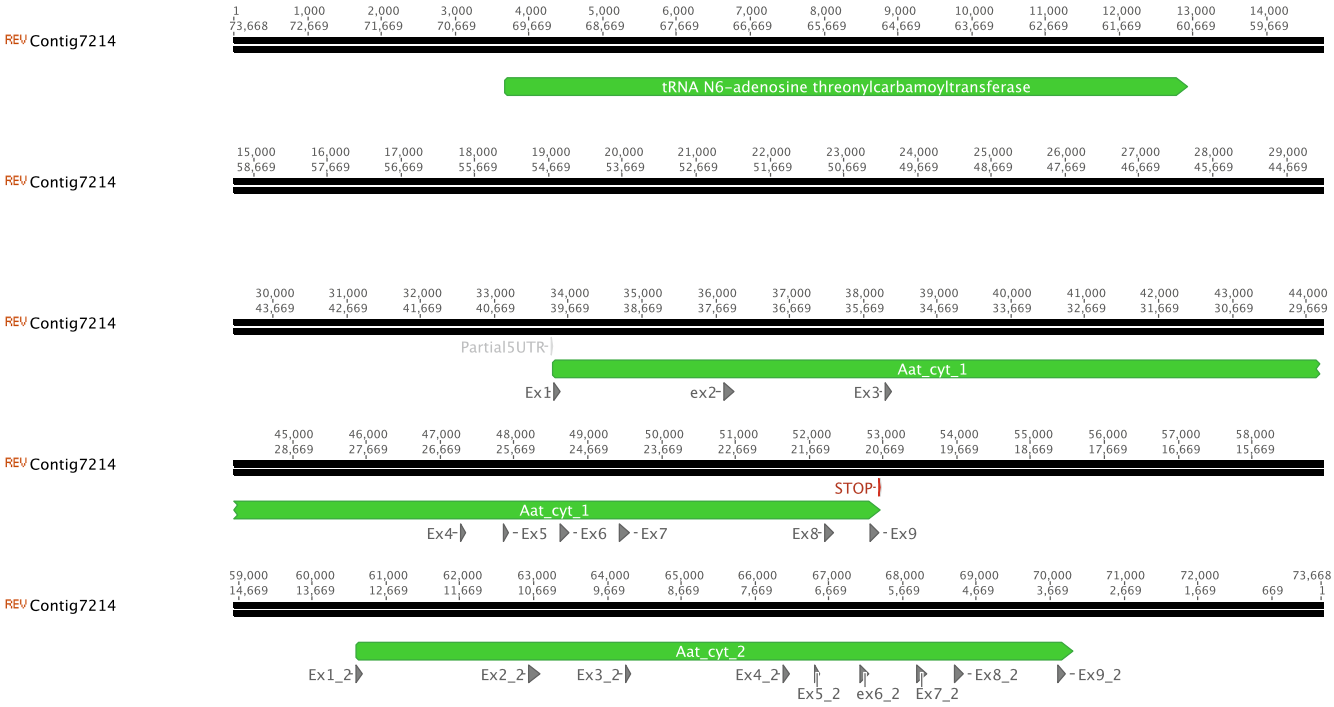

Fig. 6. *Littorina saxatilis* genome contig 7214 containing two copies of the cyt-Aat gene (green arrows). Both copies consist of 9 exons (grey arrows) and 8 introns. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4. Discussion

Several genome studies reveal intriguing patterns of genomic divergence between *L. saxatilis* from different microhabitats (Westram et al., 2018; Faria et al., 2019; Morales et al., 2019; Koch et al., 2021), and it is timely to investigate their connections to single candidate genes

reported by earlier studies of allozymes. A fruitful approach is to combine genomic studies with classical cloning and Sanger sequencing, as was done for mannose-6-phosphate isomerase in the barnacles (Nunez et al., 2020) and arginine kinase in *L. fabalis* (Le Moan et al., 2022). We envisage that the gene structure and allelic variation in aspartate aminostransferase described here will aid future genomic *Littorina* studies.

4.1. Aat allele variation

In this study we characterized the genetic variation behind strong AAT allozyme gradients previously described in *L. saxatilis*. At the DNA level, we found three different alleles (Aat^{100}, Aat^{120}-N and Aat^{120}-S) differing by 2 to 4 amino acid substitutions, explaining the somewhat confusing pattern observed at allozyme electrophoresis. The Aat^{100} allele is present across the species' range in Europe, from Iceland and Norway to Spain and Portugal (Johannesson and Johannesson, 1989; Johannesson et al., 1993; this study). The allozyme AAT^{120} is revealed to be two different variants separated by 4 substitutions, coincidentally having the same isoelectric point. One variant is present across all northern populations (Aat^{120}-N) and the other one is found only in the Iberian Peninsula (Aat^{120}-S, in Portugal and Spain). This explains why the allozyme allele clines in these populations appeared to be reversed from the clines in the northern populations (Fig. 1).

Iberian populations of many marine shallow water species are separated from the more northern populations by a phylogeographic break, and they also harbour genetically unique lineages (Roman and Palumbi, 2004; Jolly et al., 2005; Hoarau et al., 2007; Remerie et al., 2009; Campo et al., 2010). This is also true for *L. saxatilis* (Panova et al., 2011) and its relative *L. fabalis* (Sotelo et al., 2020). Here we report another unique Iberian variant, this time of a gene that was earlier pinpointed as candidate for strong differential selection, suggesting that Iberian populations may have evolved a unique genetic solution of an essentially similar problem of adaptation to desiccation. Interestingly, the high shore snails from the Northern Spanish site (Burela) had both the Northern and Southern Aat^{120}-alleles. The Burela population has been previously shown to share genetic variation with northern European populations, unlike more southern sites in Spain and Portugal (Doellman et al., 2011; Butlin et al., 2014) and is possibly located very close to the break. Other phylogeographically distinct *L. saxatilis* populations from the Atlantic coasts of Canada and US, as well as the isolated populations in Venice and South Africa, also show allozyme polymorphism with AAT^{100} and AAT^{120} alleles (Sundberg et al., 1990). While the former is likely to be the same allele everywhere, sequencing is needed to find out whether the Northern or Southern variant of the AAT^{120} allele is present in these populations.

Littorina arcana and *L. compressa*, sister-species to *L. saxatilis*, share the allozyme alleles with the latter species. (Although the direct sequencing was performed for very few individuals for these two species, the allele frequencies were confirmed in the pool-seq data). *Littorina compressa* is a low shore species (Reid, 1996) and we only found the Aat^{100} allele, which agrees with previous allozyme data (Knight and Ward, 1991). In contrast, the vertical range of *L. arcana* overlaps with *L. saxatilis* and in some sites even extends higher on the shore (Reid, 1996). Accordingly, we found both Aat^{100} and Aat^{120}-N alleles in this species, also in agreement with earlier allozyme studies (Sundberg et al., 1990). Allele sharing between these three species is most likely ancestral polymorphism since they have been shown to share genetic variation in both mitochondrial and nuclear genomes (Stankowski et al., 2020; Maltseva et al., 2022a) and to be very similar at the proteome level (Maltseva et al., 2020). Thus, the polymorphism present in AAT is ancestral to both *L. arcana* and *L. saxatilis* and in both species maintained by divergent selection between high and low shore microhabitats.

The reconstructed phylogeny of the different alleles showed two distinct clades: one with Aat^{120}-N haplotypes and the other containing Aat^{100}, Aat^{120}-S and *L. fabalis* haplotypes. The Aat^{120}-S appears to be the most recent variant, derived from the Aat^{100} allele, probably in the Iberian populations. The separation of the Aat^{100} and Aat^{120}-N alleles is much older since it precedes the divergence of the *L. fabalis* allele. A similar pattern was shown for the arginine kinase gene in *L. fabalis,* that also has two diverged clades and one of them was shared with *L. saxatilis* (Le Moan et al., 2022). A plausible explanation of these patterns is that the ancestral polymorphism in this lineage that has been retained in some species by strong differential selection.

4.2. Gene expression vs. structural protein variation in Aat

The steep allele clines found in many regions together with the fast recovery of clines after a natural perturbation suggest strong differential selection for alternative *Aat* alleles at low and high shore levels (Johannesson and Johannesson, 1989; Johannesson et al., 1995), although we cannot reject that the target of selection is one or more loci tightly linked to the *Aat* on the chromosome. A role of the *Aat* variation in the adaptation to different shore levels is further supported by the differences in enzyme activity between the AAT allozymes (Panova and Johannesson, 2004). The activity differences can be mediated by structural variation in the protein sequences or by differential expression of the alleles (Wray, 2007; Hoekstra and Coyne, 2007; Dalziel et al., 2009). Our results reject the differences at the gene expression level in the different genotypes. Further, we did not observe up-regulation of the *Aat* during the desiccation, suggesting that the *Aat* gene is expressed constitutively, and the enzymatic flux is instead regulated by the structural differences between the allozymes, for example affecting substrate affinity.

There are several known examples when one or a few structural mutations in a protein have large effects on the phenotype and fitness, e.g., cytosolic malate dehydrogenase in the limpet genus *Lottia* (Dong and Somero, 2009), melanocortin 1 receptor in mice (Hoekstra et al., 2006), hemoglobin in the cod (Andersen et al., 2009), lactate dehydrogenase in the killifish (Powers and Schulte, 1998) and tetrodotoxin receptor in Garter snakes (Geffeney et al., 2005). Aspartate aminotransferase in *Littorina* snails may be another one, and it would be interesting to investigate the effect of this polymorphism on fitness,

The general pattern that have emerged from the allozyme studies is that all low shore species in northern Atlantic (e.g. *L. littorea*, *L. fabalis*, *L. obtusata* and *L. compressa*) are all fixed or nearly so for one AAT allele (Zaslavskaya et al., 1992; Johannesson, 1992; Tatarenkov and Johannesson, 1994), while species distributed over much of the vertical intertidal or living high up on the shore (*L. saxatilis*, *L. arcana*, and *Melarhaphe neritoides*) are strongly polymorphic in the AAT locus (Sundberg et al., 1990; Johannesson, 1992). In the gene tree, both the Aat^{100} (low shore) and Aat^{120}-N (northern high-shore alleles) appeared to be old variants, widespread across the *L. saxatilis* range and present in its sister-species, and therefore likely to be older than the species themselves. Likely selection pressures on the *Aat* in snails is the maintenance of normal metabolic rates during desiccation at the lower shore and surviving the long desiccation by lowering the metabolism at the high shore (McMahon, 1990). The Aat^{120}-S allele described here is however more recently derived from the Aat^{100} and its frequency slightly higher at the low show (Johannesson et al. 1993); it would be very interesting to compare the enzymatic activity of this new structural variant with the Aat^{100}, which is also common at the low shore.

A remaining alternative hypothesis is that the differentiation on the *Aat* gene results from selection on a linked locus. The *Littorina* genome contains many large inversions, and they are likely to be important for rapid adaptation of the snails to various environments (Faria et al., 2019; Koch et al., 2021). Morales et al. (2019) found several of these inversions to be involved in the differentiation between low and shore snails; one of them is a large inversion on linkage group 12 (26.3 cM, or almost the half of the linkage group) containing the contig 7214 with the *Aat* gene. While the differences in protein sequences and especially in the enzymatic activity of the AAT alleles, coupled to the aspartate-

malate flux and the metabolism rate in the intertidal snails, support the selection for different alleles of the *Aat* gene, we cannot exclude the possibility that selection is acting on another tightly linked locus within the inversion. The study by Morales et al. (2019) focused on the whole-genome patterns using F_{ST} calculated over 500 bp genome regions, which is longer than, for example, the *Aat* exons, and is likely to miss the divergence caused by few non-synonymous substitutions. Genomic studies of the variation between low- and high-shore snails at a finer scale are thus necessary to further investigate possible targets of selection within the inversion.

4.3. Genomic structure of the Aat gene

Using the obtained coding sequences, we reconstructed the *Aat* gene in the *L. saxatilis* genome. The structure of the cytosolic *Aat* gene has been previously characterized in detail for the Pacific oyster *Magallana (Crassostrea) gigas* (Boutet et al., 2005). In the oyster the gene had 9 exons and 8 introns, a molecular weight of 45.3 kDa and an isoelectric point of 8.15, i.e., is very similar to the predicted protein in *Littorina*.

In the *L. saxatilis* genome there is another copy of the gene, containing the same number of exons and introns. It seems however to be non-functional since it is neither recovered in our RNA sequencing nor in recently produced transcriptome assemblies for the species (accession numbers GHUM00000000.1, GHUL00000000.1 and GHPE00000000.1). Considering high divergence between the gene copies (10% in exons, > 50% of introns) and their split in the protein tree, the second gene copy appears to be a product of an old gene duplication event preceding the evolution of the *Aat* alleles and divergence of the *Littorina* species. The exact timing of this duplication event in the evolution of littorinids can be determined in the future, when the genome for other species in the family become available. In addition, there is a truncated, yet expressed gene copy, found in the transcriptome assemblies and in the 3′UTR sequencing by the GeneRacer.

Interestingly, we also found multiple gene copies of two other genes recently characterized in the *L. saxatilis* genome: arginine kinase (Le Moan et al., 2022) and the sperm protein LOSP (Maltseva et al., 2022b). Further, a comparative array hybridization analysis of the *L. saxatilis* genome indicated that up to 10% of the protein-coding genes may be present in multiple copies (Panova et al., 2014). Together these results suggest that duplication of genes and genome regions is quite common in *L. saxatilis*.

An interesting question is whether the apparently common gene duplications have some functional role in the snails and may also have contributed to their rapid evolution and on-going divergence. In case of aspartate aminotransferase and arginine kinase the second gene copies appear to be pseudo-genes - non-functional degrading sequences (this study and Le Moan et al., 2022). In case of the LOSP protein however the genome of *L. saxatilis* contains two potentially functional copies of the gene (Maltseva et al., 2022b). Moreover, the LOSP is suggested to be a so-called orphan gene – a novel gene evolved in *Littorina* lineage by duplication and re-use of exons from other genes, horizontal transfer of genome fragments from bacteria by transposable elements and expansion of repeats (Maltseva et al., 2022b). The fact that all three candidate genes characterized in the *Littorina* genome so far appear to be present in the multiple copies indicates an interesting possibility that repetitive nature of *Littorina* genome may provide novel functional advantages for the snails and may represent another, undiscovered, key genomic feature behind their rapid adaptations.

Kerstin Johannesson reports financial support was provided by Swedish research council Vetenskapsrådet.

CRediT authorship contribution statement

Felix Mittermayer: Writing – original draft, Methodology, Formal analysis, Investigation. **Cecilia Helmerson:** Methodology, Formal analysis, Investigation. **Mårten Duvetorp:** Methodology, Formal analysis, Investigation. **Kerstin Johannesson:** Conceptualization, Investigation, Writing – review & editing, Funding acquisition. **Marina Panova:** Writing – original draft, Conceptualization, Methodology, Investigation, Supervision, Project administration.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Kerstin Johannesson reports financial support was provided by Swedish research council Vetenskapsrådet.

Data availability

All data is submitted to the NCBI archives

Acknowledgements

We thank Anna-Karin Ring and Emma Johansson for their help with the laboratory work, Alexandra Sa Pinto for providing partial sequences of the *Aat* transcripts, Hernan Morales for providing bam-files from the pool-seq data and Magnus Alm Rosenblad for many helpful discussions on the *Aat* gene structure in the *L. saxatilis* genome. The work was supported by the Swedish research council Vetenskapsrådet, [grant number 2017-03798] and the Centre for Marine Evolutionary Biology at Gothenburg University.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gene.2023.147517.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215 (3), 403–410.
Andersen, Ø., 2012. Hemoglobin polymorphisms in Atlantic cod - a review of 50 years of study. Mar. Genomics 8, 59–65.
Andersen, Ø., Wetten, O.F., De Rosa, M.C., Andre, C., Carelli Alinovi, C., Colafranceschi, M., Colosimo, A., 2009. Haemoglobin polymorphisms affect the oxygen-binding properties in Atlantic cod populations. Proc. R. Soc. B Biol. Sci. 276 (1658), 833–841.
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 (10), 2044–2072.
Boutet, I., Meistertzheim, A.L., Tanguy, A., Thébault, M.T., Moraga, D., 2005. Molecular characterization and expression of the gene encoding aspartate aminotransferase from the Pacific oyster Crassostrea gigas exposed to environmental stressors. Comp. Biochem. Physiol. C: Toxicol. Pharmacol. 140 (1), 69–78.
Butlin, R.K., Saura, M., Charrier, G., Jackson, B., André, C., Caballero, A., Coyne, J.A., Galindo, J., Grahame, J.W., Hollander, J., Kemppainen, P., Martínez-Fernández, M., Panova, M., Quesada, H., Johannesson, K., Rolan-Alvarez, E., 2014. Parallel evolution of local adaptation and reproductive isolation in the face of gene flow. Evolution 68 (4), 935–949.
Campo, D., Molares, J., Garcia, L., Fernandez-Rueda, P., Garcia-Gonzalez, C., Garcia-Vazquez, E., 2010. Phylogeography of the European stalked barnacle (*Pollicipes pollicipes*): identification of glacial refugia. Mar. Biol. 157, 147–156.
Canbäck, B., André, A., Galindo, J., Johannesson, K., Johansson, T., Panova, M., Tunlid, A., Butlin, R., 2012. The *Littorina* Sequence Database (LSD) - an online resource for genomic data. Mol. Ecol. Resour. 12 (1), 142–148.
Clement, M., Snell, Q., Walker, P., Posada, D., Crandall, K., 2002. TCS: estimating gene genealogies. In: Parallel and Distributed Processing Symposium, International, vol. 3. IEEE Computer Society, p. 0184.
Comeault, A.A., Flaxman, S.M., Riesch, R., Curran, E., Soria-Carrasco, V., Gompert, Z., Nosil, P., 2015. Selection on a genetic polymorphism counteracts ecological speciation in a stick insect. Curr. Biol. 25 (15), 1975–1981.
Crawford, D.L., Powers, D.A., 1989. Molecular basis of evolutionary adaptation at the lactate dehydrogenase-B locus in the fish *Fundulus heteroclitus*. Proc. Natl. Acad. Sci. 86 (23), 9365–9369.
Dalziel, A.C., Rogers, S.M., Schulte, P.M., 2009. Linking genotypes to phenotypes and fitness: how mechanistic biology can inform molecular ecology. Mol. Ecol. 18 (24), 4997–5017.
de Zwaan, A., 1991. Molluscs. In: Bryant, C. (Ed.), Metazoan Life Without Oxygen. Chapman and Hall, London, pp. 186–217.
Dean, A.M., Thornton, J.W., 2007. Mechanistic approaches to the study of evolution: the functional synthesis. Nat. Rev. Genet. 8 (9), 675–688.

Doellman, M.M., Trussell, G.C., Grahame, J.W., Vollmer, S.V., 2011. Phylogeographic analysis reveals a deep lineage split within North Atlantic *Littorina saxatilis*. Proc. R. Soc. B Biol. Sci. 278 (1722), 3175–3183.

Dong, Y., Somero, G.N., 2009. Temperature adaptation of cytosolic malate dehydrogenases of limpets (genus *Lottia*): differences in stability and function due to minor changes in sequence correlate with biogeographic and vertical distributions. J. Exp. Biol. 212 (2), 169–177.

Faria, R., Chaube, P., Morales, H., Larsson, T., Lemmon, A.R., Lemmon, E.M., Rafajlovic, M., Panova, M., Ravinet, M., Johannesson, K., Westram, A.M., Butlin, R. K., 2019. Multiple chromosomal rearrangements in a hybrid zone between *Littorina saxatilis* ecotypes. Mol. Ecol. 28 (6), 1375–1393.

Geffeney, S.L., Fujimoto, E., Brodie III, E.D., Brodie Jr, E.D., Ruben, P.C., 2005. Evolutionary diversification of TTX-resistant sodium channels in a predator–prey interaction. Nature 434 (7034), 759–763.

Hedrick, P.W., Ginevan, M.E., Ewing, E.P., 1976. Genetic polymorphism in heterogeneous environments. Annu. Rev. Ecol. Syst. 7 (1), 1–32.

Hoarau, G., Coyer, J.A., Veldsink, J.H., Stam, W.T., Olsen, J.L., 2007. Glacial refugia and recolonization pathways in the brown seaweed *Fucus serratus*. Mol. Ecol. 16 (17), 3606–3616.

Hochachka, P.W., 1980. Living Without Oxygen: Closed and Opensystems in Hypoxia Tolerance. Harward University Press, Cambridge, Mass., USA, pp. 27–41.

Hoekstra, H.E., Coyne, J.A., 2007. The locus of evolution: evo devo and the genetics of adaptation. Evolution 61 (5), 995–1016.

Hoekstra, H.E., Hirschmann, R.J., Bundey, R.A., Insel, P.A., Crossland, J.P., 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. Science 313 (5783), 101–104.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17 (8), 754–755.

Hummel, H., Sommer, A., Bogaards, R.H., Pörtner, H.O., 1997. Variation in genetic traits of the lugworm *Arenicola marina*: temperature related expression of mitochondrial allozymes? Mar. Ecol. Prog. Ser. 159, 189–195.

Janson, K., Ward, R.D., 1984. Microgeographic variation in allozyme and shell characters in *Littorina saxatilis* Olivi (Prosobranchia: Littorinidae). Biol. J. Linn. Soc. 22 (4), 289–307.

Johannesson, K., Johannesson, B., Rolán-Alvarez, E., 1993. Morphological differentiation and genetic cohesiveness over a microenvironmental gradient in the marine snail *Littorina saxatilis*. Evolution 47 (6), 1770–1787.

Johannesson, K., Tatarenkov, A., 1997. Allozyme variation in a snail (*Littorina saxatilis*) - deconfounding the effects of microhabitat and gene flow. Evolution 51 (2), 402–409.

Johannesson, K., Butlin, R.K., Panova, M., Westram, A.M., 2017. Mechanisms of adaptive divergence and speciation in *Littorina saxatilis*: integrating knowledge from ecology and genetics with new data emerging from genomic studies. In: Oleksiak, M.F., Rajora, P. (Eds.), Population Genomics: Marine Organisms. Springer.

Johannesson, K., Johannesson, B., Lundgren, U., 1995. Strong natural selection causes microscale allozyme variation in a marine snail. Proc. Natl. Acad. Sci. 92(7), 2602–2606.

Johannesson, K., Johannesson, B., 1989. Differences in allele frequencies of *Aat* between high- and mid-rocky shore populations of *Littorina saxatilis* (Olivi) suggest selection in this enzyme locus. Genet. Res. 54 (1), 7–11.

Johannesson, K., 1992. Genetic variability and large scale differentiation in two species of littorinid gastropods with planktotrophic development, *Littorina littorea* (L.) and *Melarhaphe (Littorina) neritoides* (L.) (Prosobranchia: Littorinacea), with notes on a mass occurence of *M. neritoides* in Sweden. Biol. J. Linnean Soc 47(3), 285–299.

Jolly, M.T., Jollivet, D., Gentil, F., Thiébaut, É., Viard, F., 2005. Sharp genetic break between Atlantic and English Channel populations of the polychaete *Pectinaria koreni*, along the North coast of France. Heredity 94 (1), 23–32.

Kent, W.J., 2002. BLAT—The BLAST-like alignment tool. Genome Res. 12 (4), 656–664.

Knight, A.J., Ward, R.D., 1991. The genetic relationships of three taxa in the *Littorina saxatilis* species complex (Prosobranchia: Littorinidae). J. Moll. Stud. 57 (1), 81–91.

Koch, E.L., Morales, H., Larsson, J., Westram, A.M., Faria, R., Lemmon, A.R., Lemmon, E. M., Johannesson, K., Butlin, R.K., 2021. Genetic variation for adaptive traits is associated with polymorphic inversions in *Littorina saxatilis*. Evol. Lett. 5 (3), 196–213.

Koch, E.L., Ravinet, M., Westram, A.M., Johannesson, K., Butlin, R.K., 2022. Genetic architecture of repeated phenotypic divergence in *Littorina saxatilis* ecotype evolution. Evolution 76 (10), 2332–2346.

Koehn, R.K., Zera, A.J., Hall, J.G., 1983. Enzyme polymorphism and natural selection. In: Nei, M., Koehn, R.K. (Eds.), Evolution of Genes and Proteins. Sinauer, Sunderland, Mass, USA, pp. 115–136.

Laurent, S., Pfeifer, S.P., Settles, M., Hunter, S.S., Hardwick, K.M., Ormond, L., Rosenblum, E.B., 2016. The population genomics of rapid adaptation: disentangling signatures of selection and demography in white sands lizards. Mol. Ecol. 25 (1), 306–323.

Le Moan, A., Panova, M., De Jode, A., Ortega, O., Duvetorp, M., Faria, R., Butlin, R., Johannesson, K., 2022. An allozyme polymorphism is associated with a large chromosomal inversion in the marine snail *Littorina fabalis*. Evol. Appl. 16, 279–292.

Maltseva, A.L., Varfolomeeva, M.A., Lobov, A.A., Tikanova, P.O., Panova, M., Mikhailova, N.A., Granovitch, A.I., 2020. Proteomic similarity of Littorinid snails in the evolutionary context. PeerJ 8, e8546.

Maltseva, A.L., Panova, M.A.Z., Varfolomeeva, M.A., Vikhreva, D.V., Smutin, D.V., Pavlova, P.A., Maslakov, G.P., Danilov, L.G., Mikhailova, N.A., Granovitch, A.I., 2022a. Phylogeography of the closely related *Littorina (Neritrema)* species in the North-East Atlantic. Invertebr. Zool. 19 (4), 404–424.

Maltseva, A.L., Lobov, A.A., Pavlova, P.A., Panova, M., Gafarova, E.R., Marques, J., Danilov, L.G., Granovitch, A.I., 2022b. Orphan gene in *Littorina*: an unexpected role of symbionts in the host evolution. Gene 824, 146389.

Martínez-Fernández, M., Bernatchez, L., Rolán-Alvarez, E., Quesada, H., 2010. Insights into the role of differential gene expression on the ecological adaptation of the snail *Littorina saxatilis*. BMC Evol. Biol. 10 (1), 1–14.

McDonald, J.H., Kreitman, M., 1991. Adaptive protein evolution at the Adh locus in *Drosophila*. Nature 351 (6328), 652–654.

McMahon, R.F., 1990. Thermal tolerance, evaporative water loss, air-water oxygen consumption and zonation of intertidal prosobranchs: a new synthesis. In: Progress in Littorinid and Muricid Biology: Proceedings of the Second European Meeting on Littorinid Biology, Tjärnö Marine Biological Laboratory, Sweden, July 4-8, Springer, Netherlands, 1988, pp. 241–260.

Mojica, J.P., Lee, Y.W., Willis, J.H., Kelly, J.K., 2012. Spatially and temporally varying selection on intrapopulation quantitative trait loci for a life history trade-off in *Mimulus guttatus*. Mol. Ecol. 21 (15), 3718–3728.

Morales, H., Faria, R., Johannesson, K., Larsson, T., Panova, M., Westram, A.M., Butlin, R.K., 2019. Genomic architecture of parallel ecological divergence: beyond a single environmental contrast. Sci. Adv. 5, eaav9963.

Nosil, P., Funk, D.J., Ortiz-Barrientos, D., 2009. Divergent selection and heterogeneous genomic divergence. Mol. Ecol. 18 (3), 375–402.

Nunez, J.C., Flight, P.A., Neil, K.B., Rong, S., Eriksson, L.A., Ferranti, D.A., Rand, D.M., 2020. Footprints of natural selection at the mannose-6-phosphate isomerase locus in barnacles. Proc. Natl. Acad. Sci. 117 (10), 5376–5385.

Panova, M., Blakeslee, A.M.H., Miller, A.W., Mäkinen, T., Ruiz, G.M., Johannesson, K., André, C., 2011. Glacial history of the North Atlantic marine snail, *Littorina saxatilis*, inferred from distribution of mitochondrial DNA lineages. PLoS One 6 (3), e17511.

Panova, M., Johannesson, K., 2004. Microscale variation in *Aat* (aspartate aminotransferase) is supported by activity differences between upper and lower shore allozymes of *Littorina saxatilis*. Mar. Biol. 144, 1157–1164.

Panova, M., Johansson, T., Canbäck, B., Bentzer, J., Alm, R.M., Johannesson, K., Tunlid, A., André, C., 2014. Species and gene divergence in *Littorina* snails detected by array comparative genomic hybridization. BMC Genom. 15, 687.

Powers, D.A., Schulte, P.M., 1998. Evolutionary adaptations of gene structure and expression in natural populations in relation to a changing environment: a multidisciplinary approach to address the million-year saga of a small fish. J. Exp. Zool. 282 (1–2), 71–94.

Ravinet, M., 2018. Notes from a snail island: littorinid evolution and adaptation. Mol. Ecol. 27 (13), 2781–2789.

Ravinet, M., Faria, R., Butlin, R.K., Galindo, J., Bierne, N., Rafajlović, M., Westram, A.M., 2017. Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. J. Evol. Biol. 30 (8), 1450–1477.

Reid, D.G., 1996. Systematics and Evolution of *Littorina*. Ray Society, London.

Reid, D.G., Dyal, P., Williams, S.T., 2012. A global molecular phylogeny of 147 periwinkle species (Gastropoda, Littorininae). Zool. Scr. 41 (2), 125–136.

Remerie, T., Vierstraete, A., Weekers, P.H., Vanfleteren, J.R., Vanreusel, A., 2009. Phylogeography of an estuarine mysid, *Neomysis integer* (Crustacea, Mysida), along the north-east Atlantic coasts. J. Biogeogr. 36 (1), 39–54.

Roman, J.O.E., Palumbi, S.R., 2004. A global invader at home: population structure of the green crab, *Carcinus maenas*, Europe. Mol. Ecol. 13 (10), 2891–2898.

Slater, G.S.C., Birney, E., 2005. Automated generation of heuristics for biological sequence comparison. BMC Bioinf. 6 (1), 31.

Sokolova, I.M., Pörtner, H.O., 2001. Physiological adaptations to high intertidal life involve improved water conservation abilities and metabolic rate depression in *Littorina saxatilis*. Mar. Ecol. Prog. Ser. 224, 171–186.

Sotelo, G., Duvetorp, M., Costa, D., Panova, M., Johannesson, K., Faria, R., 2020. Phylogeographic history of flat periwinkles, *Littorina fabalis* and *L. obtusata*. BMC Evol. Biol. 20, 1–18.

Stankowski, S., Westram, A.M., Zagrodzka, Z.B., Eyres, I., Broquet, T., Johannesson, K., Butlin, R.K., 2020. The evolution of strong reproductive isolation between sympatric intertidal snails. Philos. Trans. Roy. Soc. B (Biology) 375 (1806), 20190545.

Storz, J.F., Wheat, C.W., 2010. Integrating evolutionary and functional approaches to infer adaptation at specific loci. Evolution 64 (9), 2489–2509.

Sundberg, P., Knight, A., Ward, R.D., Johannesson, K., 1990. Estimating the phylogeny in *Littorina saxatilis* Olivi from enzyme data: methodological considerations. Hydrobiologia 193, 29–40.

Tatarenkov, A., Johannesson, K., 1994. Habitat related allozyme variation on a microgeographical scale in the marine snail *Littorina mariae* (Prosobranchia: Littorinacea). Biol. J. Linn. Soc. 53 (2), 105–125.

Tigano, A., Friesen, V.L., 2016. Genomics of local adaptation with gene flow. Mol. Ecol. 25 (10), 2144–2164.

Watt, W.B., 1977. Adaptation at specific loci. I. Natural selection on phosphoglucose isomerase of *Colias* butterflies: biochemical and population aspects. Genetics 87 (1), 177–194.

Westram, A.M., Rafajlovíc, M., Chaube, P., Faria, R., Larsson, T., Panova, M., Ravinet, M., Blomberg, A., Mehlig, B., Johannesson, K., Butlin, R., 2018. Clines on the seashore: the genomic architecture underlying rapid divergence in the face of gene flow. Evol. Lett. 2, 297–309.

Wheat, C.W., Watt, W.B., Pollock, D.D., Schulte, P.M., 2006. From DNA to fitness differences: sequences and structures of adaptive variants of *Colias* phosphoglucose isomerase (PGI). Mol. Biol. Evol. 23 (3), 499–512.

Winefield, C.S., Farnden, K.J.F., Reynolds, P.H.S., Marshall, C.J., 1995. Evolutionary analysis of aspartate aminotransferases. J. Mol. Evol. 40, 455–463.

Wray, G.A., 2007. The evolutionary significance of cis-regulatory mutations. Nat. Rev. Genet. 8 (3), 206–216.

Zaslavskaya, N.I., Sergievsky, S.O., Tatarenkov, A.N., 1992. Allozyme similarity of Atlantic and Pacific species of *Littorina* (Gastropoda: Littorinidae). J. Moll. Stud. 58 (4), 377–384.