Transcriptome Expression of Biomineralization Genes in *Littoraria flava* Gastropod in Brazilian Rocky Shore Reveals Evidence of Local Adaptation

Camilla A. Santos [iD][1],*, Gabriel G. Sonoda[1], Thainá Cortez[1], Luiz L. Coutinho[2], and Sónia C.S. Andrade[1],*

[1]Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, São Paulo, SP, Brazil

[2]Departamento de Ciência Animal, Escola Superior de Agricultura Luiz de Queiroz (ESALQ), Universidade de São Paulo, Piracicaba, São Paulo, SP, Brazil

*Corresponding authors: E-mails: camillasantos@usp.br; soniacsandrade@ib.usp.br.

Abstract

Understanding how selection shapes population differentiation and local adaptation in marine species remains one of the greatest challenges in the field of evolutionary biology. The selection of genes in response to environment-specific factors and microenvironmental variation often results in chaotic genetic patchiness, which is commonly observed in rocky shore organisms. To identify these genes, the expression profile of the marine gastropod *Littoraria flava* collected from four Southeast Brazilian locations in ten rocky shore sites was analyzed. In this first *L. flava* transcriptome, 250,641 unigenes were generated, and 24% returned hits after functional annotation. Independent paired comparisons between 1) transects, 2) sites within transects, and 3) sites from different transects were performed for differential expression, detecting 8,622 unique differentially expressed genes. Araçá (AR) and São João (SJ) transect comparisons showed the most divergent gene products. For local adaptation, fitness-related differentially expressed genes were chosen for selection tests. Nine and 24 genes under adaptative and purifying selection, respectively, were most related to biomineralization in AR and chaperones in SJ. The biomineralization-genes perlucin and gigasin-6 were positively selected exclusively in the site toward the open ocean in AR, with sequence variants leading to pronounced protein structure changes. Despite an intense gene flow among *L. flava* populations due to its planktonic larva, gene expression patterns within transects may be the result of selective pressures. Our findings represent the first step in understanding how microenvironmental genetic variation is maintained in rocky shore populations and the mechanisms underlying local adaptation in marine species.

Key words: RNA-seq, environmental heterogeneity, gene differential expression, adaptative selection, shell mineralization, *L. flava*.

Significance

The genetic structure of *Littoraria flava* was greater on a microgeographical scale, despite its planktotrophic larvae with high dispersal rates. Differential expression profiles showed that local adaptation may occur on a microgeographical scale in biomineralization genes identified under positive selection. Our results are a starting point for many marine species in understanding how population differentiation is maintained in the rocky shore populations, probably by local adaptation due to the great spatial heterogeneity of their habitat.

Introduction

The mechanisms by which gene flow, selection, and biotic and abiotic variables shape the local adaptation of a species are not yet fully understood and remain one of the greatest challenges in the field of evolutionary biology. Differences in species development and larval dispersal, along with the diversity of specific microhabitats, such as rocky shores, are crucial to many population parameters, including colonization success, sexual rate, and genetic structure (Haye et al. 2014; Fernández et al. 2015). The intertidal rocky shore is known as one of the most heterogeneous environments ever described, representing a transitional ecosystem between land and marine systems enabling the establishment of diverse ecotypes in many microhabitats. This environment presents a mosaic of microenvironments with unique properties, such as substratum type, topographic complexity, cracks, vegetation coverage, wave action, temperature, and species interaction, each of which shapes the adaptation and biodiversity of the communities living within them (Menge et al. 1985; Bustamante and Branch 1996; Banks and Skilleter 2007). The marine gastropod Littoraria flava is abundant in Brazilian rocky shore habitats, with a high dispersal capability due its planktotrophic larval phase and a distribution spanning from Florida (United States) to the South region of Brazil coastline. Species in this environment are typically observed gathered in small groups, depending on the tide and degree of predation (Reid 1985; Moutinho and Alves-Costa 2000; Reid et al. 2010).

Molluscs are vital for the functioning of aquatic and terrestrial ecosystems, in addition to being excellent ecological indicators of how species adapt and cope with the stresses caused by climate change (Sousa et al. 2018). This group is at the top of the list of species affected by the warming and acidification of oceans, which directly affects the calcification of their shells (Chandra Rajan and Vengatesen 2020). Owing to their wide geographic distribution and large population, many aquatic and grazers gastropods are used as model organisms in ecological disturbance studies (Wit and Palumbi 2013; Gleason 2019). The gradients of wave action and temperature in rocky shores directly influence the structure of the living communities (Bustamante and Branch 1996; Sanford and Kelly 2011).

However, few data are available in the literature about the relationship among population structure, genetic variability, and gene expression levels in rocky shore species. The use of next-generation sequencing techniques in genetic adaptation studies has provided an opportunity to gain a greater understanding of the genetic basis of local adaptation in non-model species with no genome described. This has included inferring possible differentially expressed genes (DEGs) involved in the adaptation of organisms to environmental stresses and climate change, such as rising water temperatures and exposure to heavy metals (Meyer et al. 2011; Wit and Palumbi 2013; Gleason 2019). A wider clarification of the genomic regions and candidate loci that are responsive to ecological, morphological, physiological, and developmental traits, as well as their respective local adaptation patterns, remains a huge challenge to evolutionary biology. Transcriptomics allows for connections to be made between the gene expression levels and local adaptation and is greatly advantageous for species with little genomic data available, such as gastropods. Although studies investigating differential gene expression and local adaptation in marine invertebrates are still scarce, they are completely viable and able to achieve valid results (Gleason and Burton 2015; Wit and Palumbi 2013).

In the marine gastropod L. flava, its genetic structure was described as being greater on a microgeographical scale, that is, within the same rocky shore than between large distances (Andrade et al. 2003; Andrade and Solferini 2007). However, functional analysis is necessary to infer whether local adaptation processes are responsible for this structuring. In a previous study, Andrade and Solferini (2007) analyzed L. flava populations from different rocky shore sites in three locations from the southeastern region of Brazil, one of which was also investigated in the present study. They reported a patchier structuring of L. flava on a microgeographical scale than on a large scale, despite its developmental type. Cases in which the genetic structure cannot be assigned to any geographical or temporal patterns are known as chaotic genetic patchiness (CGP) (Broquet et al. 2013). Microscale genetic structuring in rocky shore zone has been previously reported in a few species, including the crab Pachygrapsus marmoratus (Iannucci et al. 2020) and the limpet Siphonaria sp. (Johnson and Black 1982). However, attempts to explain the CGP in rocky shore populations are not widely found, despite the fact that invertebrates that inhabit intertidal habitats are particularly vulnerable to environmental fluctuations, and species must adapt to extreme abiotic variations (Johannesson and Tatarenkov 1997; Wilkins 1977; Schmidt and Rand 1999, 2001). In this study, we attempted to address the lack of the use of a functional approach for the study of rocky shore species by verifying whether this pattern was maintained in L. flava when considering its gene expression levels and local adaptation. To this end, we explored the gene expression profiles in different L. flava populations from Brazilian rocky shores. More specifically, we evaluated whether the gene expression profiles differed among the distinct sites within the same rocky shore and among different localities. Among the thousands of DEGs identified, we focused on those related to fitness traits, such as shell calcification and oxidative stress. By comparing large numbers of DEGs and diverse gene products, we aimed to verify whether the diversity in gene expression signatures among locations at various distances from the sea was related to local adaptation. Although many genes under selection were identified, biomineralization genes were positively selected at particular locations, indicating that adaptation processes occur on a microscale.

Results

Transcriptome Data Overview

We generated a total of 767,812,038 reads, of which 445,176,839 (58%) were mapped against the reference constructed by Trinity (supplementary table S1, Supplementary Material online). A total of 60,109 (24%) out of 250,641 unigenes presented hits and were used in the analysis of DEGs. From this total, 1) 57,658, 2) 45,187, 3) 46,918, 4) 26,632, and 5) 19,281 unigenes returned hits in the nr-NCBI, SwissProt, Kyoto Encyclopedia of Genes and Genomes (KEGG), Gene Ontology (GO), and COG databases, respectively, with 4,771 genes showing simultaneous hits in all five databases (Heberle et al. 2015; supplementary fig. S1, Supplementary Material online). Among over 60,000 genes with hits, 1,084 (2%) showed a high similarity with mollusc species. All the genes with BlastX hits were used as a background reference for expression analysis. Independent and paired comparisons were performed in the following groups: 1) between different transects, 2) between different sites within the same transect, and 3) between corresponding site distances from different transects. In total, 8,622 unique DEGs (false discovery rate [FDR] < 0.05) were identified, considering all comparisons together (table 1 and fig. 1). The genes upregulated for the first location were automatically downregulated for the second, and vice versa. Regarding all the DEG comparisons performed in this study, we focused our results and discussion in Araçá (AR) and São João (SJ) locations due to their evident distinct expression profiles and particular classes of fitness DEGs identified, biomineralization for AR and oxidative stress to SJ.

DEGs identified by Comparison of Different Transects

We performed six comparisons between the four sampled transects and identified up to 4,212 DEGs between AR and SJ (fig. 1A and supplementary table S2, Supplementary Material online). The gene expression profiles between these samples were clearly distinct and completely separated in different clades represented by Euclidian distance heatmaps (fig. 2A). Figure 3 shows how much the expression of the genes, such as PIF, gigasin-6, perlwapin, pervitellin-22, and alfa crystallin, varied among the different transects. In the following, we report on the main upregulated genes between AR × SJ.

Araçá × São João

Among the DEGs identified, AR presented diverse gene products, highlighting genes related to 1) biomineralization (gigasin-6, PIF, and perlwapin), 2) reproduction (temptin), and 3) substrate fixation (substrate foot variant protein 1, *fp-1*), among upregulated genes (fig. 4A and supplementary table S2, Supplementary Material online). Conversely, SJ was associated with upregulated genes related to innate immunity and oxidative stress, such as peroxidase, dermatopontin, lectin,

apoptosis inhibitor, chymotrypsin, and ß-1,3-glucan-binding (fig. 4A). The GO terms identified for these locations varied widely (fig. 5 and supplementary fig. S2, Supplementary Material online). In AR, biological processes (BPs) (cell adhesion, apoptosis, and movement of cilia and microtubules), molecular functions (MFs) (actin and calmodulin binding), and cellular components (CCs) (line Z and cellular junction) are among the most frequent. On the other hand, we identified distinct GO terms in SJ, including immune-related GO terms, in addition to responses to estrogen and mercury as BP, heparanase and ß-1,3-glucan binding as MF, and lysosome as CC. Regarding the analysis of functional enrichment GO terms, 303 terms were found to be enriched for AR and 177 for SJ (supplementary table S3, Supplementary Material online). After GO enrichment, BPs related to localization, regulation of response to stimulus and locomotion and CCs as the cytoskeletal and cell projection parts were among the main terms enriched with DEGs for AR, whereas immune effector process (BP) and lytic vacuole (CC) were among those enriched for SJ.

DEGs between Different Sites within Transect

Expression profiles between two sites within the same transect were also evaluated, the comparisons of which are presented in table 1. The DE comparisons with the most informative genes detected were AR4 × AR16 (1,080 DEGs) and SJ0 × SJ32 (79 DEGs) (fig. 1B and supplementary table S2, Supplementary Material online). Although the comparison between SJ0 × SJ32 showed only 79 DEGs, the gene products identified showed expression patterns completely distinct between the sites in *L. flava* individuals, with the site samples fully separated in different clades (fig. 2B). Figure 3 shows the variation in the expression of several genes, such as perivitelline-22 and alfa crystallin, between the different transect points in AR and SJ.

Araçá (AR4 × AR16)

Littoraria flava individuals from AR4 showed high levels of expression for genes related to 1) response to several abiotic stresses, such as hypoxia, salinity, and pH (arginine kinase); 2) immunity (peptidoglycan-recognition protein and dermatopontin); and 3) foot fixation to the substrate (*fp-1*). At site 16, genes involved in 1) biomineralization as perlucin and carbonic anhydrase 1, 2) reproduction (perivitelin-2), 3) sexual behavior (temptin), and 4) defense (heavy-metal-binding HIP, glycine, glutamate, and proline-rich) were among the most upregulated (fig. 4B). The GO terms identified between these two AR sites were quite diverse (supplementary fig. S3, Supplementary Material online). The organization of the sarcomere, muscle contraction, and locomotion as BP and Z disk as CC were among the most common terms for AR4. At the AR16 site, the scenario was different when compared to AR4,

Table 1

Scheme of Paired Comparisons Performed for Differential Gene Expression Analysis

Location	Paired Comparison	Total DEGs Number	Upregulated DEGs Number per Location
Between different locations			
Araçá (AR)	Araçá × São João	4,212	2,474 AR (58%)
Praia Dura (PD)			1,738 SJ (42%)
Anchieta (AC)	Araçá × Praia Dura	1,332	339 PD (25%)
São João (SJ)			993 AR (75%)
	Araçá × Anchieta	92	40 AR (34%)
			52 AC (66%)
	Praia Dura × Anchieta	67	32 AC (38%)
			35 PD (52%)
	Praia Dura × São João	1,386	581 SJ (42%)
			805 PD (58%)
	Anchieta × São João	1,253	812 AC (65%)
			441 SJ (35%)
Between two different points from the same location			
Araçá	*P4 × P16	1,080	211 P4 (19%)
			869 P16 (81%)
	P4 × P64	246	143 P4 (58%)
			103 P64 (42%)
	P16 × P64	140	92 P16 (65%)
			48 P64 (35%)
Praia Dura	P4 × P16	52	22 P4 (42%)
			30 P16 (58%)
	P4 × P64	170	40 P4 (23%)
			130 P64 (77%)
	P16 × P64	378	99 P16 (26%)
			199 P64 (74%)
Anchieta	P4 × P64	106	70 P4 (66%)
			36 P64 (44%)
São João	P0 × P32	79	28 (35%) P0
			51 (65%) P32
Between points from different locations			
	Toward the open ocean		
Araçá (AR)	P64 AR × P32 SJ	665	396 (59%) P64 AR
Praia Dura (PD)			269 (41%) P32 SJ
Anchieta (AC)	P64 AR × P64 PD	243	115 (47%) P64 AR
São João (SJ)			128 (53%) P64 PD
	P64 AR × P64 AC	67	32 (48%) P64 AR
			35 (52%) P64 AC
	P64 AC × P32 SJ	241	113 (47%) P64 AC
			128 (53%) P32 SJ
	P64 PD × P64 AC	185	68 (36%) P64 PD
			117 (64%) P64 AC
	P64 PD × P32 SJ	179	105 (58%) P64 PD
			74 (42%) P32 SJ
	Close to shore		
	P4 AR × P0 SJ	1,338	593 (44%) P4 AR
			745 (66%) P0 SJ
	P4 PD × P0 SJ	1,245	401 (32%) P4 PD
			844 (68%) P0 SJ
	P4 AC × P0 SJ	373	204 (54%) P4 AC
			169 (46%) P0 SJ
	P4 PD × P4 AC	150	104 (69%) P4 PD
			46 (31%) P4 AC

GBE

Table 1 Continued

Location	Paired Comparison	Total DEGs Number	Upregulated DEGs Number per Location
	P4 AR × P4 AC	107	61 (57%) P4 AR
			46 (43%) P4 AC
	P4 AR × P4 PD	76	57 (75%) P4AR
			19 (25%) P4 PD

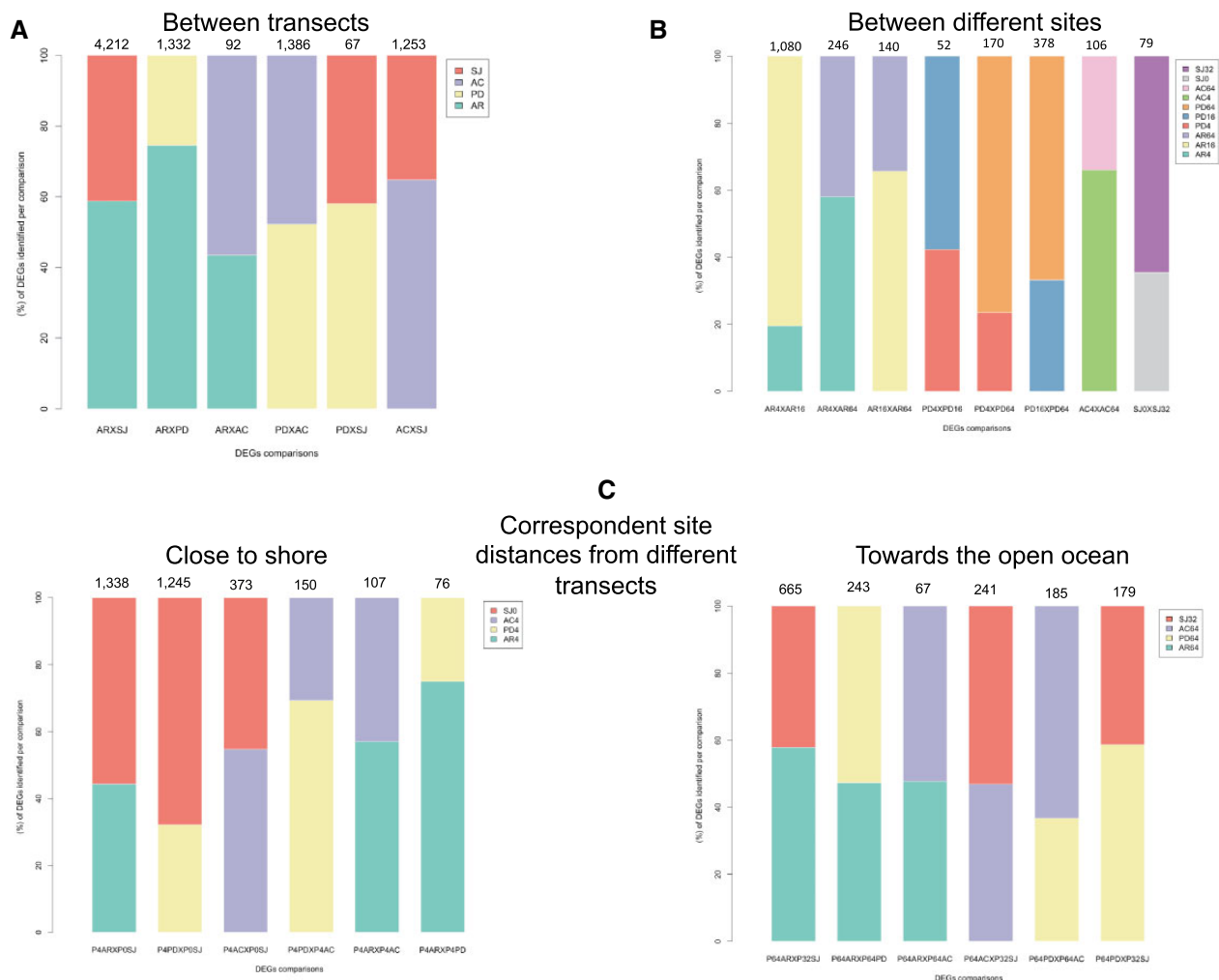NOTE.—*P, point from a given locality in the sampled location and its distance from the sea.

FIG. 1.—Bar charts representing all the differential expression comparisons performed between (A) transects, (B) different sites from the same transect, and (C) correspondent site distances from different transects, closer to shore and toward the open ocean. The comparisons performed are shown in the x axis and the percentage of upregulated genes identified in each comparison in y axis. The total number of differentially expressed genes detected for each comparison are in the top of the bars.

with cell matrix adherence, movement based on microtubules, responses to nutrients and mercury, and innate immunity as BP; lysozyme activity as MF and lysosome as CC. The GO enrichment analysis revealed 35 enriched terms for AR4 and 135 for AR16 (supplementary table S3, Supplementary Material online), with muscle-related GO terms enriched in AR4 and immunity and stress to external stimulus terms in AR16. The carbonic anhydrase biomineralization gene and perivitellin-2 gene, which are involved in embryo nutrition, were induced exclusively in AR16.
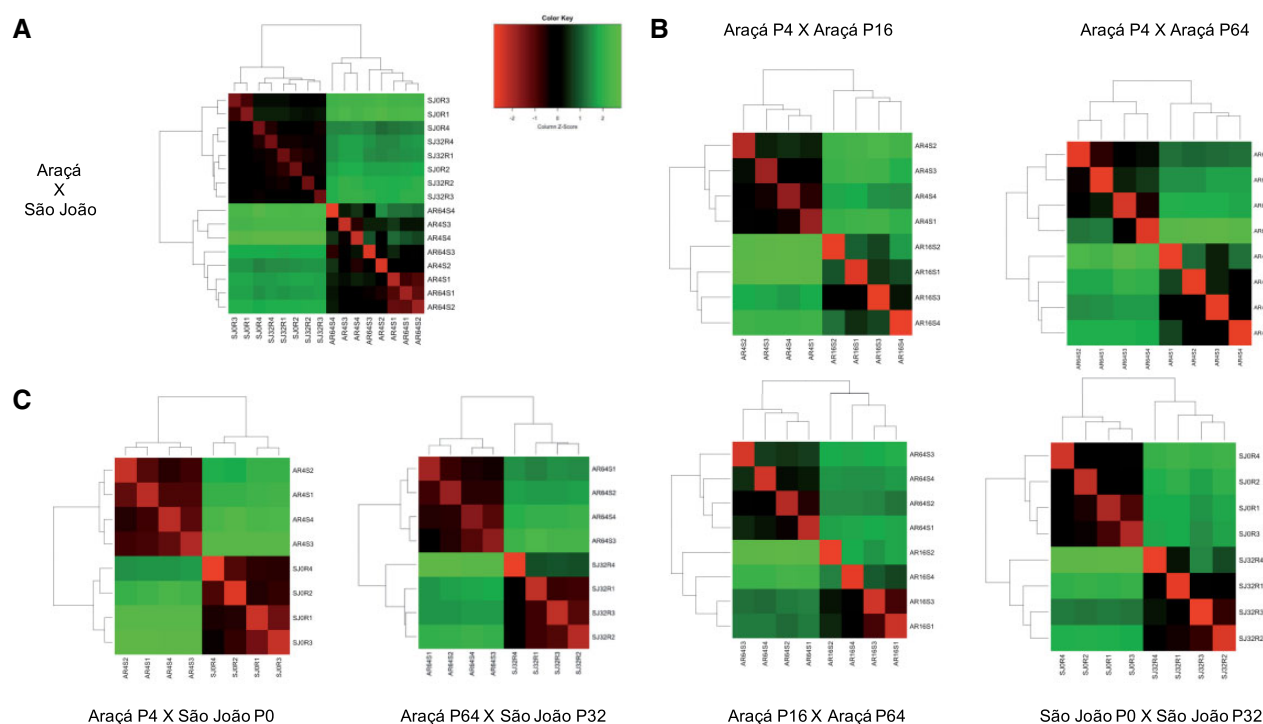
5

Fig. 2.—Euclidean distance heatmaps for the comparisons between Araçá (AR) and São João (SJ). In each comparison, samples are completely separated into different clades. Green quadrants show great genetic dissimilarity between the samples of each comparison. The numbers after the name of the locations refer to the distance in meters of the transect toward the sea (0, 4, 16, 32, and 64). The acronyms S1, S2, S3, S4, R1, R2, R3, and R4 are the biological replicates. Comparisons were performed between (A) locations, (B) within the same transect, and (C) correspondent site distances from different transects, closer to shore and toward the open ocean.

São João (SJ0 × SJ32)

Littoraria flava individuals from SJ0 present the CREBR gene, which is related to starvation and hypoxia, highly expressed (log2FC > 20). In contrast, in the SJ32 site, there was a variety and abundance of heat shock proteins (HSPs) (supplementary table S2, Supplementary Material online, and fig. 4B). When considering the GO functional annotation, the most frequent BPs for SJ0 were apoptotic processes and responses to cold and starvation, whereas for SJ32, the BPs were those related to responses to external stimuli, such as hypoxia, hydrogen peroxide, fungi, bacteria, and ultraviolet light, in addition to immune and inflammatory responses (supplementary table S3, Supplementary Material online). In terms of GO term enrichment, no terms were enriched for SJ0, whereas eight were identified for SJ32, including response to gamma radiation (supplementary table S3, Supplementary Material online).

DEGs between Correspondent Site Distances from Different Transects

Lastly, we observed the upregulated genes between the corresponding site distances from different transects. The number of DEGs varied from 665 to 67 for the sites toward the open ocean and from 1,338 to 76 for those close to shore

(fig. 1C): 1) AR4 × SJ0 with 1,338 upregulated genes, and 2) AR64 × SJ32 with 179 DEGs (table 1 and fig. 1C). The expression levels of PIF, perlucin, perivitelline-22, HSP70, and alfa crystallin genes varying between the corresponding distance sites of different transects are shown in figure 3.

P4 Araçá × P0 São João (AR4 × SJ0)

Littoraria flava individuals from AR4 and SJ0 showed completely distinct expression profiles when compared (fig. 2C). In AR4, foot protein 1 was found highly upregulated. Conversely, SJ0 presented immunity- and stress-related genes, such as chymotrypsin, glycine, glutamate, and proline-rich and HIP, in addition to the sexual behavior temptin, among the most highly induced genes (fig. 4C). The functional annotation in GO presented a greater diversity of terms for the BP. In AR4, the main BPs were sarcomere organization, nervous system development, and muscle contraction. On the other hand, in SJ0, more were identified for the stress response to nutrients, mercury, pH, estrogen, and drugs (supplementary fig. S4A, Supplementary Material online). The functional enrichment of GO terms identified 153 enriched terms for AR4 and 111 for SJ0 (supplementary table S3, Supplementary Material online).
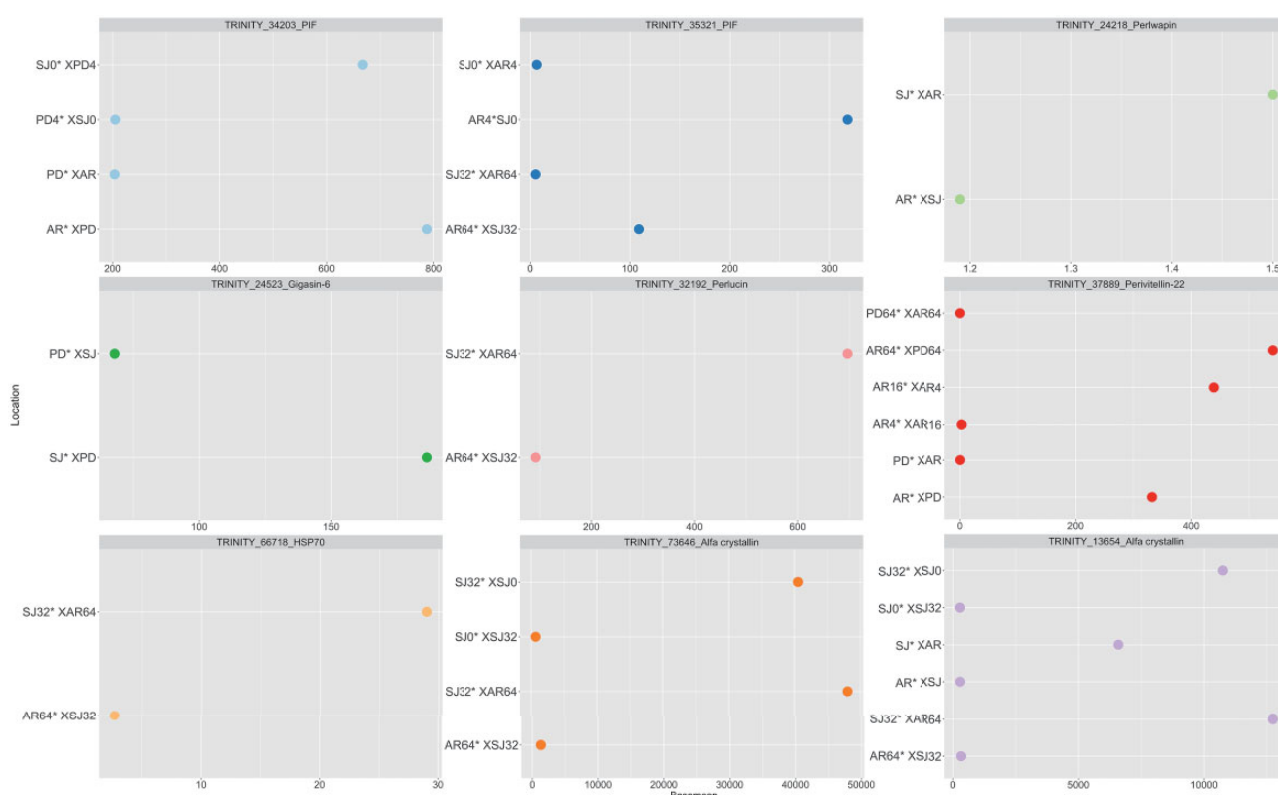
Fɪɢ. 3.—Variation in gene expression levels detected in the nine genes under positive selection discussed. The *x* axis shows the base mean values (mean of normalized gene counts) for the genes in each comparison. The *y* axis shows the comparisons where the differential expression was detected: Araçá (AR), São João (SJ), Praia Dura (PD), Araçá site 16 (AR16), Araçá site 64 (AR64), Araçá site 4 (AR4), São João site 0 (SJ0), Praia Dura site 4 (PD4), Praia Dura site 64 (PD64), and São João site 32 (SJ32). The site where the differential gene expression was detected is followed by * in the comparisons. Each rectangle represents one DEG.

P64 Araçá × P32 São João (AR64 × SJ32)

Littoraria flava individuals from AR64 showed DEGs mostly involved in biomineralization, such as PIF and perlucin, and oxidative stress. In contrast, SJ32 gastropods showed an elevated number of immune genes upregulated in the HSPs70 family (fig. 4C). The AR64 and SJ32 sites showed populations with a clear difference in their gene expression profiles (fig. 2C). In AR, the majority of BPs were related to muscle contraction and maintenance, whereas in SJ BPs, responses to pH and hypoxia were observed (supplementary fig. S4B, Supplementary Material online). We identified 54 and 75 GO enriched terms for AR and SJ, respectively (supplementary table S3, Supplementary Material online). Many HSPs were found to be upregulated exclusively in SJ32, showing a picture of intense stressful disturbances in the São João transect, more evident in sites closer to the sea.

Gene Set Enrichment Analysis

Gene set enrichment analyses (GSEAs) were carried out seeking for genes with similar expression patterns within each transect site (FDR < 0.25). Contrasting expression patterns can be observed between samples toward the open ocean and close to shore when considered all the DE comparisons performed. Among the genes that contribute the most for the enrichment are those related to actin (TRINITY2 177499 DPYL3 DANRE and TRINITY 93156 ACT3 LYTPI), perlwapin (TRINITY 24219 PWAP HALLA), and arginine kinase (TRINITY 130964 KARG HALMK) in the sites closer to shore and immunity-related trypsin and ankyrin genes in P64 (supplementary fig. S5A, Supplementary Material online). In parallel, when we consider only the samples from each transect, we see a particular set of genes contributing the most for the difference in expression levels observed in animals from different ocean distances. In AR4 we detected genes related to movement and muscle formation, such as actin (TRINITY 41424 ACT2 LUMTE) and foot protein 1 (TRINITY2 172979 FP1V1 PERVI), whereas in AR64, we highlight the perlucin gene (TRINITY 53769 PLC HALLA). In São João transect, the difference in expression profiles between samples close to shore and toward the open ocean are evident, with genes as temptin (TRINITY 35043 TEMPT APLCA) in SJ0 and many
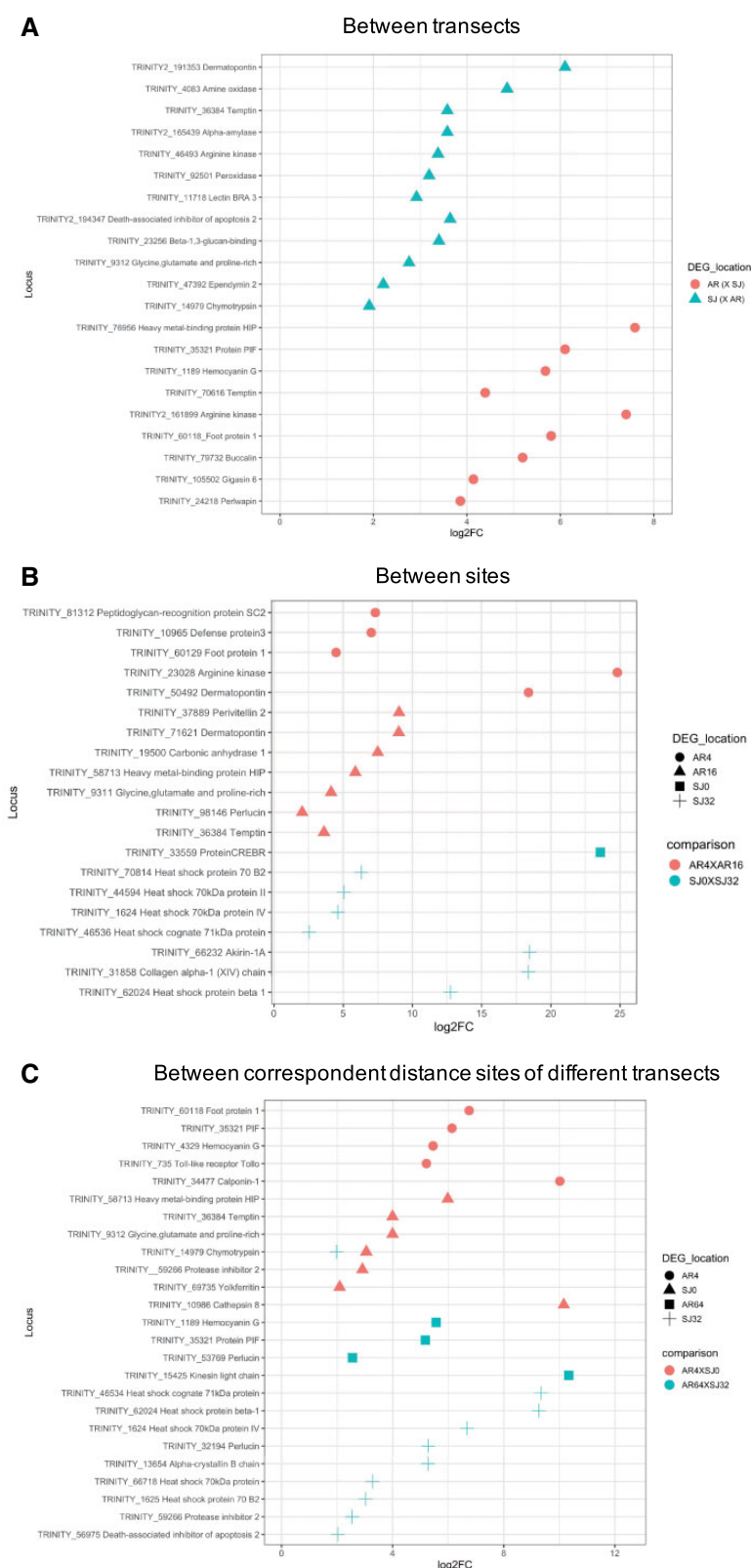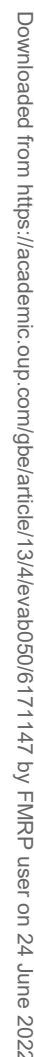
Fig. 4.—The main upregulated genes identified in Araçá and São João comparisons. (A) Between locations (AR × SJ), (B) within the same transect (AR4 × AR16 and SJ0 × SJ32), (C) correspondent site distances from different transects, close to shore (AR4 × SJ0) and toward the open ocean (AR64 × SJ32). The x axis is showing the log2FC expression values and y axis the gene names.

Fig. 5.—The most frequent GO terms and KEGG pathways identified for Araçá and São João transects, considering all the Differential Expression (DE) comparisons together. The green color refers to Araçá and red to São João. The circles' size and color are proportional to the number of terms identified, bigger and darker circles represent more counts. (*A*) Araçá GO terms, (*B*) Araçá pathways with the greatest number of KEGG Orthology (KOs) identified, (*C*) São João GO terms, and (*D*) São João pathways with the greatest number of KOs identified.

chaperone-related genes, such as HSPs (TRINITY 46534 HSP7C PONAB, TRINITY 67290 HSP7C CRIGR, TRINITY 46536 HSP7C PONAB, TRINITY 1624 HSP74 PARLI, TRINITY 62024 HSPB1 POELU, TRINITY 44594 HSP72 PARLI, and TRINITY 89501 HSP12 MEDSA) and alpha-crystallin (TRINITY 73646 CRYAB SHEEP, TRINITY 13654 CRYAB SPAJD, TRINITY2 16472 CRYAB CHICK, and TRINITY2 161696 CRYAB CHICK) in SJ32. Anchieta transect shows carbonic anhydrase (TRINITY 47 CAH1 GORGO) in the site close to the shore and biomineralization genes as PIF (TRINITY 84564 PIF PINMG and TRINITY 35321 PIF PINFU) and perlucin (TRINITY 32192 PLCL MYTGA) in AC64 among the enriched ones. At last, in Praia Dura alpha-crystallin gene (TRINITY 62022 CRYAB CHICK) was detected in PD4, whereas shell matrix MAM and LDL-receptor (TRINITY 129554 MLRP2 ACRMI) gene was observed in PD64 (supplementary fig. S5*B*, Supplementary Material online).

Araçá and São João GO and KEGG Hits

The GO and KEGG hits for the upregulated genes between the AR and SJ transects, including all the comparisons performed, showed a marked difference between their expression profiles (fig. 5). In the AR transect, muscle movement, organization, and maintenance terms are abundant, such as BP sarcomere organization, muscle contraction, muscle development, and locomotory behavior. Endocytosis, regulation of the actin cytoskeleton, MAPK signaling, and focal adhesion were among the pathways with more KOs (KEGG Orthology) identified (fig. 5). Conversely, a scenario of intense oxidative stress was observed in SJ, with lysosomes as the CC and response to estrogen, nutrient, stimulus, pH, drugs, hypoxia, and immune responses as BP. Intense oxidative stress was also observed in the São João transect, with pathways related to lysosomes, oxidative phosphorylation, and the peroxisome with the greatest number of KOs (fig. 5).

Molecular Evolution Analyses

The transcriptome assemblies for the 13 gastropod species and *L. flava* diversity and completeness are provided in supplementary figure S6, Supplementary Material online. After filtering, the 29 orthogroups kept had only the sequences data from the seven species of Caenogastropoda (*Hemifusus tuba, Batillaria attramentaria, Echinolittorina malaccana, Littorina obtusata, Littorina fabalis, Littorina littorea,*

and *Littorina saxatilis*) and *L. flava* from the ten sites sampled. These orthogroups were aligned for the selection tests. In the site (codon) test, 98 codons were identified under positive selection ($P > 0.95$ and $\omega > 1$) distributed in nine genes: lectin 3, perlucin, hemocyanin G, perlwapin, gigasin-6, foot protein variant 1, arginine kinase, and two heavy-metal-binding protein HIP loci (supplementary table S4, Supplementary Material online). We focused only on the selection results of AR and SJ transects because of the aforementioned contrasting scenario between them. The lectin 3 protein presents six sites under positive selection, all of which were located in the C-type lectin domain, with a substitution of a valine (V) to an aspartate acid (D) at position 54 in the amino acid sequences of the PD4, AR4, and PD64 sites. The perlwapin protein, involved in shell biomineralization, contained 48 codons under positive selection, three of which were located in the whey acid protein (WAP) domain, an important region for protein function (supplementary table S4, Supplementary Material online). The perlucin protein, also involved in shell formation, contains 36 sites under positive selection, ten of which contained amino acid substitutions by to an acid aspartic residue (D) in the AR64 sequence only, in addition to the substitution of a serine residue (S) to a glutamine (Q) in the 170 position, in the C-type lectin domain (supplementary table S4, Supplementary Material online). Another biomineralization protein with sites under positive selection is gigasin-6, in which eight out of 12 sites under positive selection are different in the AR64 sequence. To verify whether the perlucin and gigasin-6 genes were under positive selection in the AR64 site, we performed a branch-site selection test, where the same sites for both proteins were detected positively selected. The foot protein variant 1, responsible for the adherence of the gastropod foot in substrate and mucus production, was found to have two out of 18 sites under positive selection in the AR64 site. The arginine kinase protein presented two patterns from amino acid positions 48 to 92, in which SJ0, SJ32, AR4, and AR16 shared one pattern and AR64 presented another pattern. Despite the fact that only residues 48, 52, 74, 82, and 92 were found to be under positive selection, the whole region between 48 and 92 residues was located in the phosphagen kinase N domain (supplementary table S4, Supplementary Material online). The two heavy-metal-binding HIP orthogroups also presented codons under positive selection (supplementary table S4, Supplementary Material online).

Regarding branch analysis, only genes under purifying selection were identified ($P < 0.05$ and $\omega < 1$). Twenty-four out of 29 identified orthogroups were found to be under purifying selection in specific locations (supplementary table S5, Supplementary Material online). Two loci were negatively selected only in the transect sites closer to the sea: perivitellin-2 (AR64, AC64, and SJ32) and gigasin-6 (AC64, PD64, and SJ32). When analyzing the conformation of gigasin-6, a structural change was observed in the protein main chain at position 216 in AR64, where an isoleucine residue (I) was replaced by an aspartic acid (D), resulting in the replacement of a hydrophobic and nonpolar residue for an acidic, hydrophilic, and negatively charged residue (supplementary figs. S7 and S8, Supplementary Material online). In addition, when the opposite substitution was performed (D > I), an expansion of the cavity volume of approximately 92 Å³ was observed. The sequence of the perlucin protein identified in AR64 showed two residue substitutions, resulting in alterations in the protein conformation compared with the other transect locations, 46 S > D and 49 N > T. The 46 substitutions involved the replacement of a neutral serine residue (S) by a negative aspartic acid (D), whereas the 49 substitutions involved the replacement of the target residue (directly acting with other molecules): Asparagine (N) was exposed and threonine (T) was buried. In contrast, ten amino acids were replaced by aspartic acid (D) in a row in the AR64 perlucin protein chain. These substitutions primarily involved neutral, nonpolar, and hydrophobic residues being replaced by an acidic, polar, negatively charged, and hydrophilic one (fig. 6 and supplementary fig. S8, Supplementary Material online), potentially resulting in an alteration of the interaction of the protein with the target.

At the codon site tests, the most frequent GO terms identified in the genes under positive selection were those related to peptidase inhibitor and creatine kinase activities, in addition to the KEGG KO K23636, which was also related to peptidases and inhibitors. On the other hand, for the genes detected under purifying selection in the branch test, the most common GO terms were related to chitin metabolism, responses to hypoxia, gamma radiation, hydrogen peroxide, mostly at site AR64, and lysozyme (SJ0) and pheromone (AR64) activities. KO K09542, corresponding to HSPs, was also identified (supplementary fig. S9, Supplementary Material online).

Discussion

This study is an attempt to elucidate whether microscale structuring identified previously (Andrade and Solferini 2007) in *L. flava* populations from Brazilian rocky shore is related to differential gene expression and local adaptation. The periwinkles analyzed in this study inhabit rocky shores, one of their natural habitats, and one of the most heterogeneous environments known, being subjected to abrupt variation in environmental conditions. The large variety of microhabitats spreading naturally along rocky shores may lead to adaptative features exclusive from particular sites due to a strong selection that overcomes gene flow. Several studies investigating rocky shore species using molecular markers have identified clusters not belonging to their native populations, where interpopulation divergence is not explained by other factors, but not geography (Ravinet et al. 2016; Westram et al. 2018; Iannucci et al. 2020). Unexpected genotypic frequencies were
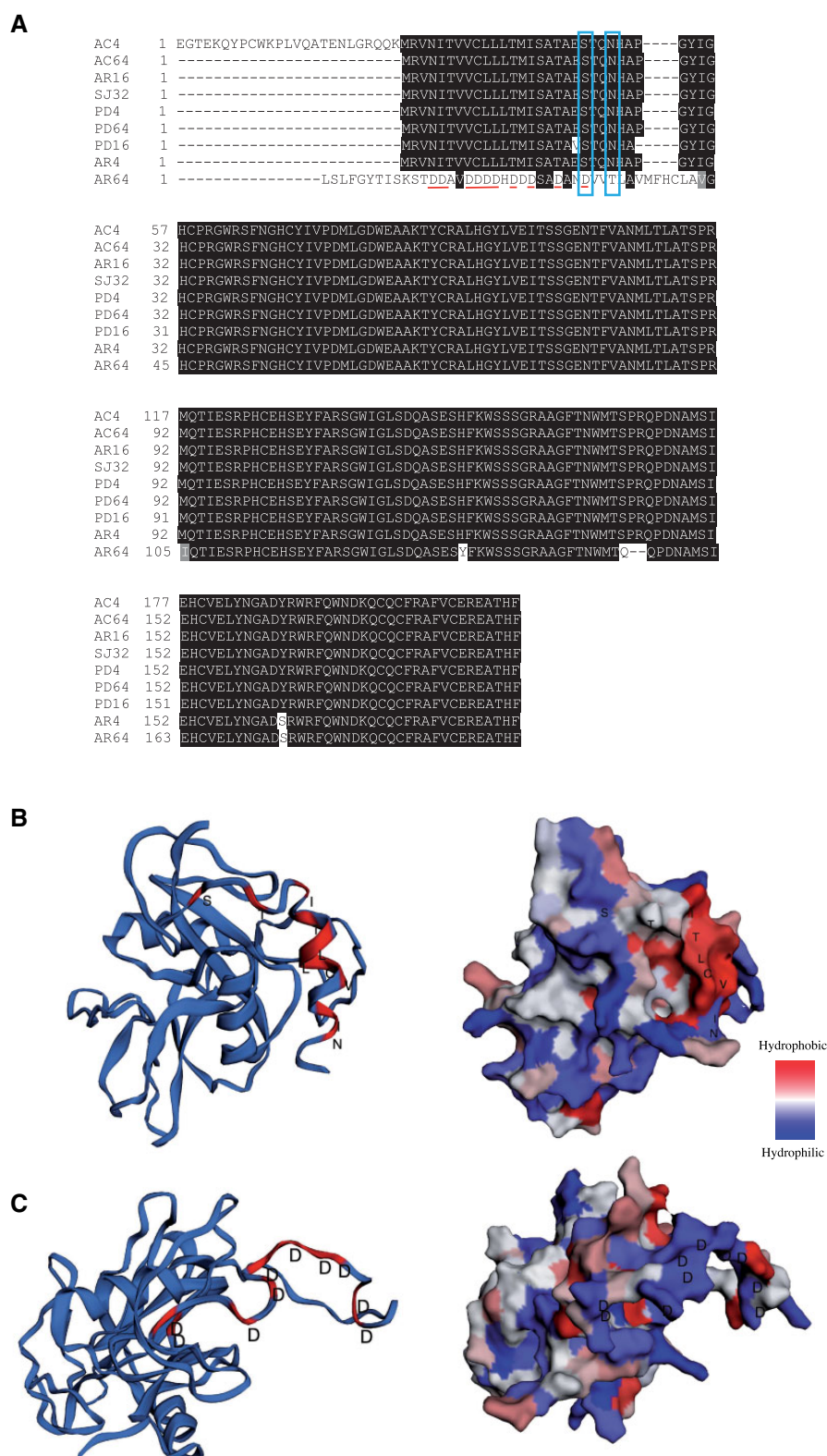
Fig. 6.—The perlucin gene orthogroup alignment and protein model. (*A*) The protein sequence alignment for nine sites sampled. The ten amino acids aspartic acid (D) substitutions in AR64 are highlighted in red. The two residues substitutions in AR64 that led to changes in protein conformation are shown by the blue squares. (*B*) The perlucin protein model based on the amino acid sequences of all the sites except AR64. The amino acids are shown by their

also observed in the limpet of the genus *Siphonaria* in 50 m along the rocky shore, which did not follow a consistent pattern. In their report, the authors suggested that the CGP observed was likely due to temporal variations in the numbers and genotypes of recruits, giving rise to a fine-scale differentiation (Johnson and Black 1982). In the present study, we used a functional approach based on gene expression levels and evidence of local adaptation in *L. flava* species as a starting point toward understanding how local heterogeneity may occur on a fine scale in rocky shore species. The aim was to provide insights that may explain the chaotic patchiness pattern observed in other marine species. Local adaptation may become even more pronounced regarding fitness-related genes that directly influence the individual's survival success. In our study, we produced the first transcriptome of the littorinid *L. flava*, a species with no genomic data available to date. Specimens were collected from ten sites in the Southeast region of Brazil coastline in four horizontal transects to the sea, allowing for an investigation of microhabitats within rocky shores. We evaluated the genetic expression profiles of *L. flava* individuals living in different locations and sites within these locations to identify their respective expression signatures and verify whether there could be a potential association between gene expression and local heterogeneity. In total, 26 differential expression comparisons were performed between different transects, sites within transects, and sites from the corresponding sea distances from different transects, revealing 8,622 unique DEGs. Several genes with products related to traits of interest and/or performance in mollusks were found to be differentially expressed after comparison, such as those related to muscle organization and maintenance, biomineralization, immunity, response to external stimuli, sexual behavior, and fixation in the substrate. Many of these genes were found to be exclusively differentially expressed in particular sites, perhaps as a response to particular microhabitats conditions in the rocky shore.

In Araçá, the main upregulated genes are involved in muscle development, biomineralization, microtubule motor activity, foot substrate adhesion, and reproduction (fig. 4). The shell of molluscs is composed of calcium carbonate crystals ($CaCO_3$), in the range of 1–5%, and the organic matrix is mainly constituted of proteins, pigments, and polysaccharides, such as chitin (Sarashina et al. 2006). The predominance of chitin in the shell matrix is in agreement with our findings, where the most frequent BP negatively selected in many locations was chitin metabolism (supplementary table S5 and fig. S9, Supplementary Material online). Calcium transporter and shell proteins are central in the formation and maintenance of shells (Furuhashi et al. 2009; Chandra Rajan and Vengatesen 2020). In our data, biomineralization genes, such as PIF, perlwapin, gigasin-6, and perlucin, were found to be upregulated, the last three of which were found to be under adaptive selection in *L. flava* individuals (fig. 4 and supplementary tables S4 and S5, Supplementary Material online).

The PIF protein family, consisting mainly of Pif 80 and Pif 97, have homologs only in the bivalves and gastropods groups and act in the regulation of nacre layer formation, working together with other shell-related proteins, such as shematrin, in shell calcification and regeneration (Lin et al. 2014). Two PIF genes were identified under purifying selection in Araçá (sites 4 and 16) and São João (sites 0 and 32). This suggests that PIF gene expression is crucial for *L. flava* individuals of AR4, given its high expression level (log2FC > 6) with a sequence conserved by purifying selection (Marek and Tomala 2018). The perlwapin gene is also related to biomineralization and has been reported to act, protect, and remodel the shell matrix from proteolytic degradation in the gastropod *Haliotis laevigata* (Treccani et al. 2006). It is rich in WAP domains, which are known to be related to shell maintenance (Addadi et al. 1987; Marin 2012; Freer et al. 2014). Our findings indicated that three codons in the perlwapin gene were under adaptive selection in WAP domain regions, playing a potential role in shell formation. In addition, the perlwapin gene was found to be upregulated in the entire Araçá transect. Although the branch test indicated that this protein was under purifying selection in the Araçá transect (supplementary table S5, Supplementary Material online), some of its codons were found to be under positive selection (supplementary table S4, Supplementary Material online). These results indicate that, even though some residues were under adaptive changes (i.e., mutations resulting in changes of residues are more prone to fixate), a great extent of the proteins were under purifying selection, masking the few sites under positive selection in the branch model. Another biomineralization gene, perlucin, was first isolated from the nacre layer of *Ha. laevigata* shell and is a carbohydrate-binding protein calcium-dependent (Freer et al. 2014; Hüning et al. 2016). The protein involved in nacre composition is responsible for its mechanical properties, such as increasing fracture strength, in addition to nucleating calcium carbonate crystals and participating in crystal formation when sodium chloride is present (Blank et al. 2003). The perlucin gene was found to be upregulated in SJ32 (fig. 4) only and was detected with an accelerated evolution rate in AR64, with a distinct amino acid

---

Fig. 6—Continued

respective letters. The left image shows the protein without the surface and the right one with the surface. Note the highly hydrophobic site in red where the amino acids are shown. (*C*) The perlucin protein model based on the amino acid sequence of AR64. The amino acids are shown by their respective letters. The left image shows the protein without the surface and the right one with the surface. Note the highly hydrophilic site in blue where the amino acids are shown. The proteins models were diagrammed by Ezmol (http://www.sbg.bio.ic.ac.uk/ezmol/, last accessed October 20, 2020).

sequence for ten codons in AR64, which led to the replacement of a hydrophobic region with a highly hydrophilic one (fig. 6). A similar scenario was observed in the gigasin-6 gene, with an eight-codon sequence exclusive of AR64 and under positive selection in this location only. The 216-position isoleucine residue (I) was replaced by an aspartic acid (D) in AR64 individuals, resulting in the replacement of a hydrophobic and nonpolar residue for a hydrophilic, acid, and negatively charged one (supplementary fig. S7, Supplementary Material online). Substitutions for aspartic acid leading to hydrophilic aspartic acid-rich shell matrix proteins are known to regulate and induce biomineralization in mytilid mussels and are determinant in the formation of many forms of $CaCO_3$ structures when the beta-chitin form is present (Hüning et al. 2016). Another feature of aspartic acid-rich regions is their capacity to act as chelators, attracting and increasing the concentration of calcium ions close to the shell matrix (Addadi et al. 1987; Marin et al. 2007). When we analyzed this replacement in the opposite direction (D > I), an expansion in the cavity region was detected. Cavities are determinants for protein function, given that they are sites at which proteins bind to target molecules, such as other proteins and metabolites (Xu et al. 2018). The gigasin-6 gene was detected with its sequence preserved by purifying selection in the most distant site from the land in SJ32 (supplementary table S5, Supplementary Material online), whereas in AR64 the new variants were positively selected (supplementary table S4, Supplementary Material online), characterizing a potential scenario of local adaptation.

Adaptative divergence among populations may occur from a few meters to many kilometers, mirroring the intensity of selection according to the environmental conditions and gene flow (Hedgecock 1986; Sanford and Kelly 2011). Another point acting directly in local adaptation is the heterogeneity of an environment, such as the intertidal rocky shore (Bustamante and Branch 1996; Banks and Skilleter 2007). Despite the small distance separating Araçá sites, our data show a differentiation in *L. flava* populations from AR64 for gigasin-6 and perlucin biomineralization genes, indicating a different selection gradient along the transect and a selection strong enough to lead to local adaptation (Hedgecock 1986; Andrade and Solferini 2007; Banks and Skilleter 2007) at this site. The perlucin genes are one of the most responsive ones regarding the particular microenvironments of Araçá transect, with a considerable difference in expression profiles between individuals from toward the open ocean (P64) and close to shore (P4) sites (supplementary fig. S5*B*, Supplementary Material online). A plausible possibility is selection during the settlement stage (Schmidt and Rand 2001), where all larvae can achieve the AR64 site, but only those carrying the set of variants that confer adaptative advantages at such sites may settle and develop at this location, characterizing a sort of ecological filter. The large number of biomineralization genes upregulated in Araçá may be a result of the lower pH

found in this beach (8,05) when compared with São João (8,12), according to the Bio-Oracle database (Assis et al. 2018). The lower pH found in Araçá waters may be a potential cause for gigasin-6 and perlucin genes under positive selection in AR64, likely due to the greater contact with an acidified sea. In addition, Araçá is located in a region with intense anthropic activities, such as fishing and release of untreated sewage (Amaral et al. 2010). Due to ocean acidification, more carrier ion proteins acting in biomineralization and/or a differentiated expression of shell formation genes may be required. This compensation through the level of increased expression of these genes is probably not natural, resulting in a greater metabolic cost for mollusk species (Sarashina et al. 2006; Melzner et al. 2011).

Genes related to fitness traits, such as immunity, biomineralization, and oxidative stress, are expected to evolve at higher rates, given their importance in survival and adaptation (Aguilera et al. 2014; Gleason and Burton 2015). Some marine species have been reported to show signals of local adaptation for tolerance to heat stress (Gleason and Burton 2015), response to salinity (Lamichhaney et al. 2012), and ocean current and sea temperature (Miller et al. 2019). As an example, we identified the perivitellin-2 gene identified under purifying selection in SJ32 and AR64. This gene is involved in yolk production and is required for embryo nutrition and development (Chen et al. 2010). In figure 3, we can see how much the perivitellin gene expression varies along the AR transect, with higher levels detected in AR16. In addition, the perivitellin-2 gene was significantly upregulated in AR64 (log2FC > 26) when compared with PD64, suggesting a great relevance for species survival. The *L. flava* spawning period is likely to be from December to March (Cardoso et al. 2007) and the specimens used in this study were collected during January, which may be a possible explanation for the perivitellin-22 high expression levels detected. It is known that the evolutionary rate of a protein is directly linked to its gene expression, with those showing higher expression levels being more conserved (Marek and Tomala 2018). Marek and Tomala (2018) showed that gene expression was negatively related to the level of polymorphism in functional regions of protein-coding genes in natural populations. Therefore, a possible explanation is purifying selection acting against changes in protein sequences of genes expressed at high levels, as an attempt to conserve the sequences of genes that are essential for species survival. When we consider the Araçá transect, despite the small distance between the transect sites, the purifying selective force in AR64 seems to be strong enough to increase the chance of local adaptation. The adaptative divergence in marine invertebrates may occur on a fine scale (along a few meters), in vertical or horizontal gradients in the intertidal zone areas, such as observed for *L. saxatilis* the rocky shore (Janson 1982) in response to environmental factors. On the other hand, the perivitellin-22 gene was also found under negative selection in another point toward the open

ocean, in SJ32. When we consider the Araçá and São João transects, they are more than 400 km apart from each other, and in this case, the selective gradient would not be strong enough to overcome a possible gene flow. It is expected that *L. flava* planktotrophic larvae can survive for up to 10 weeks and disperse for several kilometers (Reid 1985; Andrade et al. 2003; Reid et al. 2010), where population differentiation could be triggered by local adaptation processes (Sanford and Kelly 2011). Studies regarding phenotypic variation, such as shell thickness and body size, have reported adaptive divergence in *Littorina scutulata* and *L. obtusata* populations separated by distances greater than 400 km (Yamada 1989; Trussell 2000).

The responsive genes triggered in a disturbing environment may give us a clue regarding the unfavorable conditions that gastropods are exposed to, such as drugs, toxins, excessive heat, and insufficient oxygen. São João individuals show DEGs related to immunity and response to oxidative stress, mostly HSPs, and in toward the sea site, SJ32 (fig. 4). The HSP70 gene was under negative selection at the SJ32 site, in addition to being upregulated at this site, indicating a central protein involved in *L. flava* survival and adaptation (supplementary table S5, Supplementary Material online). Our data suggest that *L. flava* individuals in SJ32 may be more tolerant to disturbances, due to the large number of chaperones, such as HSP and alpha crystallin, among the most responsive genes at this point of the transect (supplementary fig. S5, Supplementary Material online). *Littoraria flava* is a supralittoral species that cannot stay submerged for many hours, a time that may not be enough to resist a high tide period without protein damage. The São João transect shows many GO terms related to responses to external disturbances, mostly estrogen, followed by pH, mercury, drugs, hypoxia, hypotonic salinity, and starvation, in agreement with the high number of HSPs and alpha crystallin chains genes over expressed. HSPs are known as stress proteins that work in the suitable compartmentation and folding of proteins (Hofmann 1999). They play a major role in thermotolerance but can also have their expression triggered by other types of stressor factors, such as hypoxia and toxin/drug exposure (Feder and Hofmann 1999). Oxidative stress genes appear upregulated mostly in the site closest to the water (SJ32), suggesting one or more stress factors in the São João Sea, such as drugs and strong and high tides. Once collected, the animals were readily submerged in liquid nitrogen, reducing the chances that the elevated HSP expression levels were due to the stressful factors in the collection procedure.

The frequent response to estrogen identified in *L. flava* individuals from SJ32 site may be a result of the constant and long gastropod's contact with organotin compounds, such as tributyltin (TBT) or triphenyltin (TPT) used worldwide as antifouling paints in ships, artificial structures and boats since 1960. Continuous contact with TBT leads to imposex, an endocrine disruption in which male sexual organs emerges

on female gastropods that start to behave as males (Smith 1971), already reported for gastropods from three other localities on the Rio de Janeiro coast (Fernandez et al. 2002; Cardoso et al. 2009). São João is located in a region with many fishing stations and trades, resulting in intense boat traffic and high concentrations of pollutants, including heavy metals, in the Southeastern marine zone (Oigman-Pszczol and Creed 2007; do Nascimento et al. 2018). A cytochrome P450 complex enzyme responsible for the conversion of androgenic to estrogenic steroids is inhibited by competition with TBT (Nishikawa 2006). The male sex hormone androstenedione may be aromatized by mollusks and turn out estrogen when TBT is present (Le Guellec et al. 1987), even at low concentrations (Smith 1971; Horiguchi et al. 1997). This conversion was observed in the gastropod *Hinia reticulata* when the 17ß-estradiol levels increased in all groups exposed to TBT after four months of constant contact (Nishikawa 2006). This picture is similar to the one observed in São João, likely with high estrogen levels due to the presence of TBT in the water released by boats and fishing nets.

The different gene expression profiles among transects and sites suggests that there is a particular scenario for each locality and for each microhabitat within the rocky shore, in which a different set of upregulated genes is required. Thus, the distinct ecological factors shaping the variation of *L. flava* populations in these locations are also a feature to consider in local adaptation. The environmental diversity is clearly shown by the contrasting GO terms identified between sites and transects. The particular expression patterns may be assigned to the strong spatial heterogeneity found naturally in the rocky shore (Menge et al. 1985; Bustamante and Branch 1996; Banks and Skilleter 2007). The microenvironmental changes of rocky shore may act by selecting adaptive traits and lead to diversification among the populations living on it (Janson 1987), given that species in which individuals are sedentary, such as *L. flava*, will develop particular adaptation regarding the rocky shore microhabitat they are exposed to (Selander and Kaufman 1973). Genetic heterogeneity was reported in *L. saxatilis* populations located only a few meters apart in the rocky shore (Janson and Ward 1984). Signals of divergent selection involving foot area and shell aperture and shape were also reported in *L. saxatilis* populations from adjacent rocky shore regions as mechanism to prevent wave-dislodgement (Pennec et al. 2017). Environmental gradients, such as temperature and wave action, are known to potentially influence the community structure in the rocky shore (Sanford and Kelly 2011).

In our study, we are dealing with a single species that showed genes with potential positive selection depending on the site of the rocky shore the gastropods are living in, even within the same locality. This indicates that the distance from the sea and the microenvironmental variation among rocky shores are potential factors for local adaptation in *L. flava*. Many of the upregulated or under selection genes
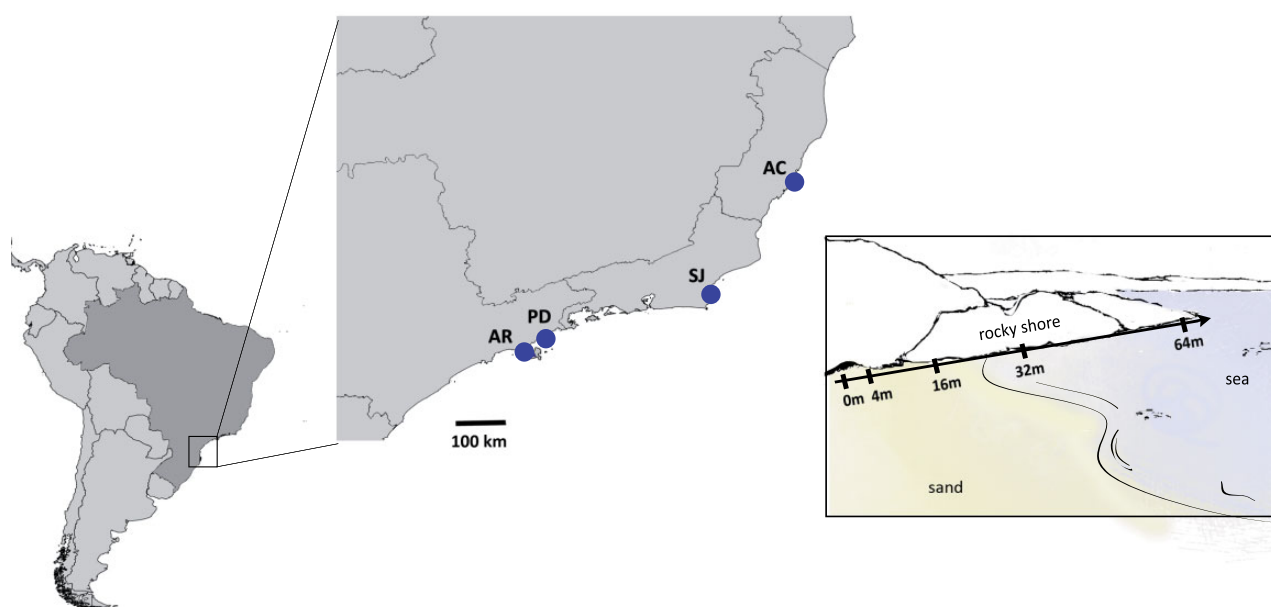
Fɪɢ. 7.—Map showing the sampling locations (blue circles) along the Southeastern region of Brazilian coastline and a transect scheme exemplifying the experimental design within each locality. Abbreviations as in table 1. The horizontal transects toward to the sea covered distances from 0 to 64 m.

identified here have never been previously described for the Littorinidae family, contributing to a more detailed understanding of biomineralization and oxidative stress physiology of the group. Our results suggest that *L. flava* populations develop particular adaptations to cope with a myriad of stress factors, such as ocean acidification, low oxygen concentrations, heat and exposure to toxins and drugs, and the responsive genes identified here may contribute to the investigation of the responses to future climate changes in marine life.

Materials and Methods

Study Area and Biological Samples

Littoraria flava specimens for RNA sequencing were collected between January and June 2018 from four localities in the Brazilian southeast region: Araçá (AR) and Praia Dura (PD) (São Paulo state), Anchieta (AC) (Espírito Santo state), and São João (SJ) (Rio de Janeiro state) (fig. 7). Araçá is a wet subtropical flat colonized by mangroves and constitutes part of Araçá Bay. This area comprises two hydrographic periods, marked by seasonal changes in temperature and salinity (Amaral et al. 2010; Dottori et al. 2015; Siegle et al. 2018). Due to its small size (~750 m wide) and its position in relation to the São Sebastião island, the tidal currents within the bay are usually weak. Sewage and oil contamination have been found in the sediment of this area (Kim et al. 2018). On the other hand, Praia Dura is formed by two rivers and has no recent records of high pollutant rates. Araçá and Praia Dura are both located in the South Brazil Bight, characterized by constant intrusions (i.e., passage of cold currents) (Fo 1979). São João is found within a well-delimited compartment

dominated by fine fluvial sediments (Muehe and Valentini 1998). The coast of Anchieta is a transitional zone located about 80 km away from Vitoria Bay, a large estuarine system (Sterza and Fernandes 2006). All sampled locations except for Praia Dura have been subjected to several anthropic activities, such as fishing, tourism, urbanization, and harbors.

Briefly, in each location, we performed horizontal transects and marked the sites using an exponential scale toward the sea to investigate different microscale distances, as described by Andrade and Solferini (2007). Specimens were collected within 1 m^2 of the marked point at sites 0, 4, 16, 32, and 64 m along the transect. Only two or three sites were sampled within each location, chosen according to the abundance of the gastropods. A total of four locations were sampled, totaling ten sites and 40 sampled individuals (table 2), with four biological replicates per site. Although the Brazilian coast possesses two different rocky shore formations, all of the sampling was performed on true rocky shores, that is, those formed by rock structures that extend from the ocean floor to above sea level. Once collected, the sampled animals were kept in liquid nitrogen until RNA isolation.

RNA Extraction and Sequencing

Total RNA was isolated from the four biological replicates of *L. flava* obtained at each site using Trizol according to Riesgo et al. (2012)'s protocol. Due to the small size of some of the specimens collected, RNA was isolated from the entire animal for all samples. The quantity was verified using a Qμbit fluorometer (Thermo Fisher Scientific) and NanoDrop spectrophotometer (Thermo Fisher Scientific). Samples with A260/280

Table 2

Details from *Littoraria flava* Samples and Collecting in Brazilian Southeastern Rocky Shores

Location/Transect	Locality (City/State) Sampling Number per Location	Latitude	Longitude	Transect Sites (Meters toward the Ocean)	Collecting Date
Araçá (AR)	São Sebastião/SP 12	23°48′47.0″S	45°24′31.3″W	AR4, AR16, and AR64	Jan/2018
Praia Dura (PD)	Ubatuba/SP 12	23°29′32.8″S	45°09′55.0″W	PD4, PD16, and PD64	Jan/2018
Anchieta (AC)	Anchieta/ES 8	20°48′37.6″S	40°39′39.7″W	AC4 and AC64	Jun/2018
São João (SJ)	Barra de São João/RJ 8	22°59′85.1″S	41°99′12.0″W	SJ0 and SJ32	Jun/2018

ratios between 1.8 and 2.2 were considered sufficiently pure. The integrity of the samples was confirmed using BioAnalyser (Agilent Technologies Inc., Santa Clara, CA), and only samples with RNA integrity number (RIN) values >6.0 were selected for subsequent analysis. A total of 40 libraries were constructed using the TruSeq RNA Library kit Preparation V2 (Illumina Inc.). The quantification step was performed using qPCR with the Library Quantification Kit Universal (KAPA Biosystems, Wilmington, MA). The libraries were grouped and run in two lanes on the Illumina HiSeq 2500 platform with 2×100 bp paired-end using the TruSeq SBS V3 kit (Illumina Inc.) at Laboratório de Biotecnologia Animal, Escola Superior de Agricultura "Luiz de Queiroz"- Universidade de São Paulo (ESALQ-USP), Piracicaba-SP, Brazil.

The quality of the raw data generated after sequencing was visualized in the FastQC version 0.10.1 software (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/, last accessed February 5, 2021). All the reads were filtered by quality in SeqyClean v.1.9.9 (https://github.com/ibest/seqy-clean, last accessed November 1, 2020) using Phred (QS) 26 and 30 for the mean and edge minimum score values and a minimum length of 65 bp. This program was also used to remove contaminant sequences from primers and vectors using the Univec database (https://www.ncbi.nlm.nih.gov/tools/vecscreen/univec/, last accessed November 1,2020).

De Novo Assembly and Functional Annotation

After filtering, de novo assembly was performed using Trinity v.2.8.4 (Grabherr et al. 2011). Twenty libraries were selected for use in the assembly, two from each of the sampled sites. A second assembly was performed with another ten samples, different from the previous ones, one from each sampled site. In order to guarantee a greater representativeness of the transcripts in the reference constructed, the two assemblies were merged into a single transcriptome after removing redundancies with >95% similarity. For assembly, 300 bp was defined as the minimum size of the contigs formed. The diversity and completeness of the generated transcriptome was assessed by predicting the ortholog genes found in the Metazoa

database (ODB 10) using BUSCO software (Seppey et al. 2019). CD-hit (Li and Godzik 2006) and DustMasker packages (Morgulis et al. 2006) were used to remove redundant contigs with >95% similarity and sequences of low complexity, respectively. The TransDecoder package (http://transdecoder.sourceforge.net/) was used to identify the candidate coding regions. Thus, functional annotation was performed in the Trinotate pipeline (https://trinotate.github.io/) using BlastX in the following databases: Uniprot (uniref90 + SwissProt) with a cut-off value of 1e10^{-3}; Gene Ontology (GO) (Ashburner et al. 2000), with GO terms: BP, MF, and CC, and KEGG (Kanehisa et al. 2012). Only genes with BlastX hits were used in the subsequent mapping stage.

Mapping and DEG Analysis

After functional annotation, the reads of the 40 *L. flava* samples were mapped against the reference unigenes with BlastX hits using Bowtie2 (v.2.3.4.3) (Langmead and Salzberg 2012). The read counts were used as input for DEG analysis. Independent comparisons between samples from 1) different locations/transects (all samples, except those from AR16 and PD16), 2) different sites within transects (all samples), and 3) the corresponding site distances from different location/transects (all samples, except those from AR16 and PD16) were performed (table 1 and supplementary fig. S10, Supplementary Material online). In comparisons (1) and (3), only samples from the transect edges (sites 0, 4, 32, and 64) could be compared due to the lack of intermediate points sampled between AC and SJ (site 16). Thus, the difference between the expression profiles from samples located at the further points of the rocky shore transects were assessed.

Differential expression analysis was performed using the R/Bioconductor DEseq2 package (Love et al. 2015). Normalization was carried out by adjusting the data distribution according to a negative binomial distribution, followed by removing the contigs with a base mean <5. The adjusted P-value for each gene was calculated using the BH method (Benjamini and Hochberg 1995), and only the genes with FDR <0.05 were considered significant for differential

expression. The GO terms enriched with DEGs were analyzed with GOSeq v.1.24.0 (Young et al. 2010) implemented in R. The categories enriched were calculated using the Wallenius approximation (Wallenius 1963) and *P*-values were adjusted using the BH method. The categories with an adjusted *P*-value (FDR) <0.05 were considered enriched. We also performed a GSEA in order to identify gene classes with common expression patterns. The analysis was carried out in GSEA software v.4.1.0 (Subramanian et al. 2005) using 1,000 permutations and FDR <0.25, accordingly to manual.

Molecular Evolution Inference of Potential Fitness-Related Genes

After the functional annotation of DEGs, our main focus was on fitness-related genes, such as biomineralization (Sokolova et al. 2012; Mackenzie et al. 2014) and oxidative stress ones (Song et al. 2014; da Silva Cantinha et al. 2017). Given their relevance for survival and as fitness-related genes, these genes had a greater chance of being under selection (Merilä and Sheldon 1999). We selected 74 DEGs with gene products related to potential adaptative traits to check for signs of selection in *L. flava* obtained from different rocky shore transects and sites. We obtained the transcriptome data of 13 gastropod species from NCBI-SRA: *Crepidula navicella* (SRR3168550 and SRR3168551), *Chrysomallon squamiferum* (SRR8599680), *E. malaccana* (SRR3217890), *Ha. laevigata* (SRR6678000), *Tritonia diomedea* (SRR2004329), *B. attramentaria* (SRR6214970), *Haliotis rufescens* (SRR8956768 and SRR8956798), *L. obtusata* (SRR9849873), *L. fabalis* (SRR9849868), *L. littorea* (SRR11015452), *L. saxatilis* (SRR9651717), *H. tuba* (ERR3077388), and *Microhedyle glandulifera* (SRR1505118) to identify DEGs under selection in *L. flava* from different rocky shores locations. We used *L. flava* gastropods from four transects and ten different sites, as indicated in table 2. Four samples from each location were assembled together using the Trinity package, as previously described. OrthoMCL v.2.1 (Fischer et al. 2011) was used to identify orthologous clusters, including the 74 DEGs and the *L. flava* sequences from the ten points. Only orthogroups containing, at least, eight *L. flava* samples were analyzed. The ratio of the number of nonsynonymous substitutions per nonsynonymous site (d*N*) to the number of synonymous substitutions per synonymous site (d*S*) omega (ω) was used to measure the evolutionary rate. The alignment files were used as input for sites (codons), branches, and branch-site test selection in ete-evol v.3.1.2 (Huerta-Cepas et al. 2016). For sites, models M1a (relaxation) and M2a (positive-selection) were tested. Regarding the branches, the foreground branch was tested for relaxation ($P > 0.05$ and $\omega = 1$) and positive selection ($P < 0.05$ and $\omega > 1$) or purifying selection ($P < 0.05$ and $\omega < 1$).

Protein Conformation Analysis

The protein sequences were queried in the Prosite database of annotated motif descriptors (Sigrist et al. 2013) for the identification of potential domains. In order to identify alterations in protein conformation, amino acid sequences were obtained from the Phyre2 server (Kelley et al. 2015) to generate a predicted model, with .pdb files as the output. The .pdb files were then used as input in the Missense 3D database (Ittisoponpisan et al. 2019), in which amino acid substitutions leading to changes in protein conformation were detected. The proteins were visualized in Missense 3D and Ezmol (Reynolds et al. 2018) databases.

Ethics Approval

The authors declare that sampling of specimens was performed in the respect of ethical rules and Brazilian laws. All samplings were authorized by the Instituto Chico Mendes de Conservação da Biodiversidade (ICMbio, n. 56726-1, 2016).

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

Data Availability

The raw data are available at NCBI Sequence Read Archieve (SRA) under the bioproject PRJNA665072.

Literature Cited

Addadi L, Moradian J, Shay E, Maroudas NG, Weiner S. 1987. A chemical model for the cooperation of sulfates and carboxylates in calcite crystal nucleation: relevance to biomineralization. Proc Natl Acad Sci U S A. 84(9):2732–2736.

Aguilera F, McDougall C, Degnan BM. 2014. Evolution of the tyrosinase gene family in bivalve molluscs: independent expansion of the mantle gene repertoire. Acta Biomater. 10(9):3855–3865.

Amaral ACZ, Migotto AE, Turra A, Schaeffer-Novelli Y. 2010. Araçá: biodiversidade, impactos e ameaças. Biota Neotrop. 10(1):219–264.

Andrade SCS, Magalhaes CA, Solferini VN. 2003. Patterns of genetic variability in Brazilian Littorinids (Mollusca): a macrogeographic approach. J Zool Syst. 41(4):249–255.

Andrade SCS, Solferini VN. 2007. Fine-scale genetic structure overrides macro-scale structure in a marine snail: nonrandom recruitment, demographic events or selection? Local-scale genetic structure in *L. flava*. Biol J Linn. 91(1):23–36.

Ashburner M, et al. 2000. Gene Ontology: tool for the unification of biology. Nat Genet. 25(1):25–29.

Assis J, et al. 2018. Bio-ORACLE v2.0: extending marine data layers for bioclimatic modelling. Global Ecol Biogeogr. 27(3):277–284.

Banks SA, Skilleter GA. 2007. The importance of incorporating fine-scale habitat data into the design of an intertidal marine reserve system. Biol Conserv. 138(12):13–29.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B. 57(1):289–300.

Blank S, et al. 2003. The nacre protein perlucin nucleates growth of calcium carbonate crystals. J Microsc. 212(Pt 3):280–291.

Broquet T, Viard F, Yearsley JM. 2013. Genetic drift and collective dispersal can result in chaotic genetic patchiness. Evolution 67(6):1660–1675.

Bustamante R, Branch G. 1996. Large scale patterns and trophic structure of southern African rocky shores: the roles of geographic variation and wave exposure. J Biogeogr. 23(3):339–351.

Cardoso RS, Caetano CHS, Cabrini TMB. 2009. Biphallia in imposexed females of marine gastropods: new record for *Nassarius vibex* from Brazil. Braz J Biol. 69(1):223–224.

Cardoso RS, Costa DS, Loureiro VF. 2007. Mating behavior of the marine snail *Littoraria flava* (Mollusca: Caenogastropoda) on a boulder shore of south-east Brazil. J Mar Biol Assoc U K. 87(4):947–952.

Chandra Rajan K, Vengatesen T. 2020. Molecular adaptation of molluscan biomineralization to high-CO2 oceans – the known and the unknown. Mar Environ Res. 155:104883.

Chen S, Armistead JS, Provost-Javier KN, Sakamoto JM, Rasgon JD. 2010. Duplication, concerted evolution and purifying selection drive the evolution of mosquito vitellogenin genes. BMC Evol Biol. 10(1):142.

da Silva Cantinha R, et al. 2017. *HSP70* expression in *Biomphalaria glabrata* snails exposed to cadmium. Ecotoxicol Environ Saf. 140:18–23.

do Nascimento MTL, et al. 2018. Determination of water quality, toxicity and estrogenic activity in a nearshore marine environment in Rio de Janeiro, Southeastern Brazil. Ecotoxicol Environ Saf. 149:197–202.

Dottori M, Siegle E, Castro BM. 2015. Hydrodynamics and water properties at the entrance of Araçá Bay, Brazil. Ocean Dyn. 65(12):1731–1741.

Feder ME, Hofmann GE. 1999. Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. Annu Rev Physiol. 61(1):243–282.

Fernandez MA, et al. 2002. Occurrence of imposex in *Thais haemastoma*: possible evidence of environmental contamination derived from organotin compounds in Rio de Janeiro and Fortaleza, Brazil. Cad Saude Publica. 18(2):463–476.

Fernández R, Lemer S, McIntyre E, Giribet G. 2015. Comparative phylogeography and population genetic structure of three widespread mollusc species in the Mediterranean and near Atlantic. Mar Ecol. 36(3):701–715.

Fischer S, et al. 2011. Using OrthoMCL to assign proteins to OrthoMCL-DB groups or to cluster proteomes into new ortholog groups. Curr Protoc Bioinformatics. 35(1):6.12.1–6.12.19.

Fo C. 1979. Wind driven currents in the channel of São Sebastião: winter. Bol Inst Oceanogr. 38(2):111–132.

Freer A, Bridgett S, Jiang J, Cusack M. 2014. Biomineral proteins from *Mytilus edulis* mantle tissue transcriptome. Mar Biotechnol (NY). 16(1):34–45.

Furuhashi T, Schwarzinger C, Miksik I, Smrz M, Beran A. 2009. Molluscan shell evolution with review of shell calcification hypothesis. Comp Biochem Physiol B Biochem Mol Biol. 154(3):351–371.

Gleason LU. 2019. Applications and future directions for population transcriptomics in marine invertebrates. Curr Mol Bio Rep. 5(3):116–127.

Gleason LU, Burton RS. 2015. RNA-seq reveals regional differences in transcriptome response to heat stress in the marine snail *Chlorostoma funebralis*. Mol Ecol. 24(3):610–627.

Grabherr MG, et al. 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat Biotechnol. 29(7):644–652.

Haye PA, et al. 2014. Phylogeographic structure in benthic marine invertebrates of the southeast Pacific coast of Chile with differing dispersal potential. PLoS One 9(2):e88613.

Heberle H, Meirelles GV, da Silva FR, Telles GP, Minghim R. 2015. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. BMC Bioinformatics 16(1):169.

Hedgecock D. 1986. Is gene flow from pelagic larval dispersal important in the adaptation and evolution of marine invertebrates? Bull Mar Sci. 39(2):550–564.

Hofmann GE. 1999. Ecologically relevant variation in induction and function of heat shock proteins in marine organisms. Integr Comp Biol. 39(6):889–900.

Horiguchi T, Shiraishi H, Shimizu M, Morita M. 1997. Effects of triphenyltin chloride and five other organotin compounds on the development of imposex in the rock shell, *Thais clavigera*. Environ Pollut. 95(1):85–91.

Huerta-Cepas J, Serra F, Bork P. 2016. ETE 3: reconstruction, analysis, and visualization of phylogenomic data. Mol Biol Evol. 33(6):1635–1638.

Hüning AK, et al. 2016. A shell regeneration assay to identify biomineralization candidate genes in mytilid mussels. Mar Genomics. 27:57–67.

Iannucci A, et al. 2020. Investigation of mechanisms underlying chaotic genetic patchiness in the intertidal marbled crab *Pachygrapsus marmoratus* (Brachyura: Grapsidae) across the Ligurian Sea. BMC Evol Biol. 20(1):108.

Ittisoponpisan S, et al. 2019. Can predicted protein 3D structures provide reliable insights into whether missense variants are disease associated? J Mol Biol. 431(11):2197–2212.

Janson K. 1982. Genetic and environmental effects on the growth rate of *Littorina saxatilis*. Mar Biol. 69(1):73–78.

Janson K. 1987. Allozyme and shell variation in two marine snails (Littorina, Prosobranchia) with different dispersal abilities. Biol J Linn Soc. 30(3):245–256.

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

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.

Johnson MS, Black R. 1982. Chaotic genetic patchiness in an intertidal limpet, *Siphonaria* sp. Mar Biol. 70(2):157–164.

Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. 2012. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res. 40(Database issue):D109–D114.

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 10(6):845–858.

Kim BSM, et al. 2018. Organic and inorganic contamination in sediments from Araçá Bay, São Sebastião, Brazil. Ocean Coast Manag. 164:42–51.

Lamichhaney S, et al. 2012. Population-scale sequencing reveals genetic differentiation due to local adaptation in Atlantic herring. Proc Natl Acad Sci U S A. 109(47):19345–19350.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods. 9(4):357–359.

Le Guellec D, et al. 1987. In vitro metabolism of androstenedione and identification of endogenous steroids in *Helix aspersa*. Gen Comp Endocrinol. 66(3):425–433.

Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22(13):1658–1659.

Lin Y, et al. 2014. Cloning and characterization of the shell matrix protein Shematrin in scallop *Chlamys farreri*. Acta Biochim Biophys Sin (Shanghai). 46(8):709–719.

Love MI, Anders S, Kim V, Huber W. 2015. RNA-Seq workflow: gene-level exploratory analysis and differential expression. F1000Res. 4:1070.

Mackenzie CL, et al. 2014. Ocean warming, more than acidification, reduces shell strength in a commercial shellfish species during food limitation. PLoS One 9(1):e86764.

Marek A, Tomala K. 2018. The contribution of purifying selection, linkage, and mutation bias to the negative correlation between gene expression and polymorphism density in yeast populations. Genome Biol Evol. 10(11):2986–2996.

Marin F, Le Roy N, Marie B. 2012. The formation and mineralization of mollusk shell. Front Biosci (Schol Ed). 4(3):1099–1125.

Marin F, Luquet G, Marie B, Medakovic D. 2007. Molluscan shell proteins: primary structure, origin, and evolution. Curr Top Dev Biol. 80:209–276.

Melzner F, et al. 2011. Food supply and seawater p CO 2 impact calcification and internal shell dissolution in the blue mussel *Mytilus edulis*. PLoS One 6(9):e24223.

Menge BA, Lubchenco J, Ashkenas LR. 1985. Diversity, heterogeneity and consumer pressure in a tropical rocky intertidal community. Oecologia 65(3):394–405.

Merilä J, Sheldon BC. 1999. Genetic architecture of fitness and nonfitness traits: empirical patterns and development of ideas. Heredity 83(2):103–109.

Meyer E, Aglyamova GV, Matz MV. 2011. Profiling gene expression responses of coral larvae (*Acropora millepora*) to elevated temperature and settlement inducers using a novel RNA-Seq procedure. Mol Ecol. 20(17):3599–3616.

Miller AD, et al. 2019. Local and regional scale habitat heterogeneity contribute to genetic adaptation in a commercially important marine mollusc (*Haliotis rubra*) from southeastern Australia. Mol Ecol. 28(12):3053–3072.

Morgulis A, Gertz EM, Schäffer AA, Agarwala RA. 2006. Fast and symmetric DUST implementation to mask low-complexity DNA sequences. J Comput Biol. 13(5):1028–1040.

Moutinho PRS, Alves-Costa CP. 2000. Shell size variation and aggregation behavior of *Littoraria flava* (Gastropoda: Littorinidae) on a southeastern Brazilian shore. The Veliger. 43:277–281.

Muehe D, Valentini E. 1998. O litoral do Estado do Rio de Janeiro: uma caracterização físico-ambiental. Rio de Janeiro (Brazil): Fundação de Estudos do Mar.

Nishikawa J. 2006. Imposex in marine gastropods may be caused by binding of organotins to retinoid X receptor. Mar Biol. 149(1):117–124.

Oigman-Pszczol SS, Creed JC. 2007. Quantification and classification of marine litter on beaches along Armação dos Búzios, Rio de Janeiro, Brazil. J Coast Res. 23(2):421–428.

Pennec GL, Butlin RK, Jonsson PR, et al. 2017. Adaptation to dislodgement risk on wave-swept rocky shores in the snail *Littorina saxatilis*. PLoS One 12(10):e0186901.

Ravinet M, et al. 2016. Shared and nonshared genomic divergence in parallel ecotypes of *Littorina saxatilis* at a local scale. Mol Ecol. 25(1):287–305.

Reid DG. 1985. Habitat and zonation patterns of *Littoraria* species (Gastropoda: Littorinidae) in Indo-Pacific mangrove forests. Zool J Linnean Soc. 26(1):39–68.

Reid DG, Dyal P, Williams ST. 2010. Global diversification of mangrove fauna: a molecular phylogeny of *Littoraria* (Gastropoda: Littorinidae). Mol Phylogenet Evol. 55(1):185–201.

Reynolds CR, Islam SA, Sternberg MJE. 2018. EzMol: a web server wizard for the rapid visualization and image production of protein and nucleic acid structures. J Mol Biol. 430(15):2244–2248.

Riesgo A, et al. 2012. Comparative description of ten transcriptomes of newly sequenced invertebrates and efficiency estimation of genomic sampling in non-model taxa. Front Zool. 9(1):33.

Sanford E, Kelly MW. 2011. Local adaptation in marine invertebrates. Annu Rev Mar Sci. 3(1):509–535.

Sarashina I, et al. 2006. Molecular evolution and functionally important structures of molluscan dermatopontin: implications for the origins of molluscan shell matrix proteins. J Mol Evol. 62(3):307–318.

Schmidt PS, Rand DM. 1999. Intertidal microhabitat and selection at Mpi: interlocus contrasts in the northern acorn barnacle, *Semibalanus Balanoides*. Evolution 53(1):135–146.

Schmidt PS, Rand DM. 2001. Adaptive maintenance of genetic polymorphism in an intertidal barnacle: habitat- and life-stage-specific survivorship of Mpi genotypes. Evolution 55(7):1336–1344.

Selander RK, Kaufman DW. 1973. Genic variability and strategies of adaptation in animals. Proc Natl Acad Sci U S A. 70(6):1875–1877.

Seppey M, Manni M, Zdobnov EM. 2019. BUSCO: assessing genome assembly and annotation completeness. In: Gene Prediction. New York: Humana. p. 227–245.

Siegle E, Dottori M, Villamarin BC. 2018. Hydrodynamics of a subtropical tidal flat: Araçá Bay, Brazil. Ocean Coast Manag. 164:4–13.

Sigrist CJA, et al. 2013. New and continuing developments at PROSITE. Nucleic Acids Res. 41(Database issue):D344–D347.

Smith BS. 1971. Sexuality in the American mud snail, *Nassarius obsoletus* Say. J Molluscan Stud. 39(5):377–378.

Sokolova IM, Frederich M, Bagwe R, Lannig G, Sukhotin AA. 2012. Energy homeostasis as an integrative tool for assessing limits of environmental stress tolerance in aquatic invertebrates. Mar Environ Res. 79:1–15.

Song H-M, et al. 2014. Molecular characteristics of the HSP70 gene and its differential expression in female and male golden apple snails (*Pomacea canaliculata*) under temperature stimulation. Cell Stress Chaperones. 19(4):579–589.

Sousa R, Delgado J, González JA, Freitas M, Henriques P. 2018. Marine snails of the genus *Phorcus*: biology and ecology of sentinel species for human impacts on the rocky shores. Biol Res Water 2018:141–147.

Sterza JM, Fernandes LL. 2006. Zooplankton community of the Vitória Bay estuarine system (Southeastern Brazil): characterization during a three-year study. Braz J Oceanogr. 54(2–3):95–105.

Subramanian A, et al. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 102(43):15545–15550.

Treccani L, Mann K, Heinemann F, Fritz M. 2006. Perlwapin, an abalone nacre protein with three four-disulfide core (whey acidic protein) domains, inhibits the growth of calcium carbonate crystals. Biophys J. 91(7):2601–2608.

Trussell GC. 2000. Phenotypic clines, plasticity, and morphological trade-offs in an intertidal snail. Evolution 54(1):151–166.

Wallenius KT. 1963. Biased sampling; the noncentral hypergeometric probability distribution, Stanford Univ Ca Applied Mathematics And Statistics Labs.

Westram AM, Rafajlović M, Chaube P, et al. 2018. Clines on the seashore: the genomic architecture underlying rapid divergence in the face of gene flow. Evol Lett. 2(4):297–309.

Wilkins NP. 1977. Genetic variability in littoral gastropods: phosphoglucose isomerase and phosphoglucomutase in *Patella vulgata* and *P. aspera*. Mar Biol. 40(2):151–155.

Wit PD, Palumbi SR. 2013. Transcriptome-wide polymorphisms of red abalone (*Haliotis rufescens*) reveal patterns of gene flow and local adaptation. Mol Ecol. 22(11):2884–2897.

Xu Y, et al. 2018. CavityPlus: a web server for protein cavity detection with pharmacophore modelling, allosteric site identification and covalent ligand binding ability prediction. Nucleic Acids Res. 46(W1):W374–W379.

Yamada SB. 1989. Are direct developers more locally adapted than planktonic developers? Mar Biol. 103(3):403–411.

Young MD, Wakefield MJ, Smyth GK, Oshlack A. 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol. 11(2):R14.

Associate editor: Piontkivska Helen