BMC Genomics

RESEARCH

Open Access

Survival of the fittest: genomic investigations of the bay scallop reveal a shift in population structure through a summer mortality event

Denis Grouzdev[1*], Emmanuelle Pales Espinosa[1], Stephen Tettelbach[2], Arnaud Tanguy[3], Isabelle Boutet[3], Harrison Tobi[2] and Bassem Allam[1*]

Abstract

Background Understanding the genetic basis of resilience in marine organisms is critical for conservation and management, particularly in the face of escalating environmental stress and disease outbreaks. The bay scallop *Argopecten irradians* is a commercially and recreationally important shellfish species found in estuarine and coastal environments of the United States from New England to the Gulf of Mexico. In New York, adult bay scallop populations have been decimated every summer since 2019 leading to the collapse of their fishery. These mortality events were associated with annual outbreaks of an undescribed apicomplexan parasite recently named Bay Scallop Marosporida (BSM) that disrupts scallop kidneys.

Results This study investigates host–pathogen interactions and assesses changes in population structure during BSM-associated mortality events. The research compared wild and aquacultured scallops used for stock enhancement in New York, revealing significant change in population structures throughout the mortality outbreak. The results underscore the selective pressures exerted by BSM infection and environmental stressors, as evidenced by shifts in genetic divergence and allele frequencies particularly in genes associated with kidney function, stress and infection response. Through a detailed genomic and population genetic approach, this research represents a unique case study highlighting the impact of disease on marine biodiversity and advances our understanding of the impact of summer mortality events on the scallop population in NY.

Conclusions This study highlights changes in the genomic structure of bay scallops during a BSM-associated mortality event. Identified mutations (such as the one in the nephrocystin-3-like gene) represent prime candidates for specific targeted investigations to link genotypes to phenotypes. By integrating genomic and epidemiological data, the research provides a basis for understanding the impact of disease on scallop biodiversity. These findings may help guide conservation strategies for sustainable fisheries in the face of environmental change and disease outbreaks.

Keywords Genomic adaptation, Marine genomics, Pathogen resistance, Bay scallop, Apicomplexan parasites, *Argopecten irradians*, Population genetics

*Correspondence:
Denis Grouzdev
denis.gruzdev@stonybrook.edu
Bassem Allam
bassem.allam@stonybrook.edu
Full list of author information is available at the end of the article

Background

The impact of marine diseases on population dynamics and the functioning of marine ecosystems is widely recognized [1]. In recent decades, there has been a marked increase in the occurrence of large-scale marine disease events, primarily as a result of human activities. These activities include the introduction of exotic hosts and/or pathogens and changes in environmental conditions due to global climate change [2]. Among the ecologically important marine organisms severely affected by disease are bivalve mollusks [3].

Mollusca is one of the major animal phyla on earth and in the oceans, with approximately 100,000 extant and 35,000 extinct species [4], making this group the second largest animal phylum after the arthropods. The phylum is divided into 7 classes [5, 6], 2 of which, the gastropods and the bivalves, contain 98 percent of the known living molluscan species [6]. Bivalves represent an important source of food and valuable goods (shells, pearls) around the world. In 2020, over 16 million tons of bivalves were produced from farming activities worldwide [7], representing a commercial value of nearly $30 billion US. In parallel with their economic value, bivalve mollusks represent ecologically important species in many coastal environments. Through their suspension-feeding processes, they control the transfer of matter and energy from the water column to the benthic environment. Large changes in bivalve biomass in the environment (e.g., mass mortalities, invasion of new environments) are often associated with significant shifts in ecosystem functions [8, 9].

The bay scallop, *Argopecten irradians*, plays an important commercial, cultural, and ecological role along the northeastern U.S. coastline [10]. This species is characterized by a remarkable variety of shell color patterns and a comparatively short life span [11, 12]. Introduced to China in the 1980s and 1990s, *A. irradians* has become the foundation of a highly successful aquaculture industry that produces approximately one million tons per year [13–15].

Bay scallops typically spawn in the summer, and fertilization occurs in the water column. The developing larvae remain in the plankton for 1 to 2 weeks [16]. Following settlement, bay scallops exhibit very rapid shell growth in their first year (10–12 mm/month) [17–19], typically reaching shell heights of 40 to > 50 mm by their first winter [18, 20]. These scallops will spawn in the following summer at an age of ~ 1 year and recruit to the fishery in the fall. However, the majority of the bay scallop population dies naturally at 18–22 months of age, resulting in effective spawning occurring in just one year [16]. *A. irradians* is an iconic species in New York, representing the official state shell. It used to support a multi-million dollar fishery industry, with annual commercial landings in the Peconic Bays reaching over 700,000 kg of meat in 2017–2018 [21]. However, the bay scallop population in New York experienced a dramatic decline beginning in the mid-1980s due to widespread mortalities caused by harmful brown tide algal blooms [22, 23]. These population crashes were followed by prolonged periods of low abundance, likely due to limited larval recruitment [24]. In response, intensive restoration initiatives were launched in 2006 to address these challenges. By reintroducing aquacultured scallops into key areas, these efforts successfully increased both larval settlement and adult population densities, leading to a significant recovery of the fishery starting around 2010 [20, 24]. Since 2019, however, the bay scallop population in New York has suffered from catastrophic and recurring summer mortality that has devastated the commercial fishery [21]. This collapse led the U.S. Department of Commerce to declare the bay scallop fishery in the Peconic Bays a fishery disaster. These mortality events were systematically associated with annual outbreaks of an undescribed apicomplexan parasite, recently dubbed Bay Scallop Marosporida (BSM), that disrupts the tissues of infected animals [25, 26].

To gain a deeper understanding of host-parasite interactions and evaluate the potential impact of disease pressure on the genetic diversity of scallops, restriction site-associated DNA sequencing (RAD-Seq) technology was utilized to analyze the genetic structure of both wild and aquacultured scallops from New York. This genomic approach enabled a detailed examination of changes in population structure before and after summer mortality events. The analysis revealed a scallop lineage that demonstrates enhanced resilience, suggesting it is better adapted to survive these recurrent mortality challenges.

Results

Field conditions, disease metrics and scallop survivorship

To investigate the impact of BSM infection in association with stressful summer conditions on the scallop population in New York and identify potential genetic features associated with resilience, we introduced wild and aquacultured scallops from Orient Harbor (the location with a relatively lower mortality rate) into Flanders Bay, which experiences high mortality rates due to parasitic infection. The study was structured around key time points: T0 representing the start of the experiment in June, T1 the end of July, T2 mid-August, T3 mid-September, and T4 the end of the study period in October (Fig. 1).

Environmental conditions (i.e., temperature and dissolved oxygen) in the deployment site are shown in Fig. 1a. Temperature peaked at 29°C in late July (range 15 to 29) and dissolved oxygen (range 2.5 to 10.5 mg/L) was

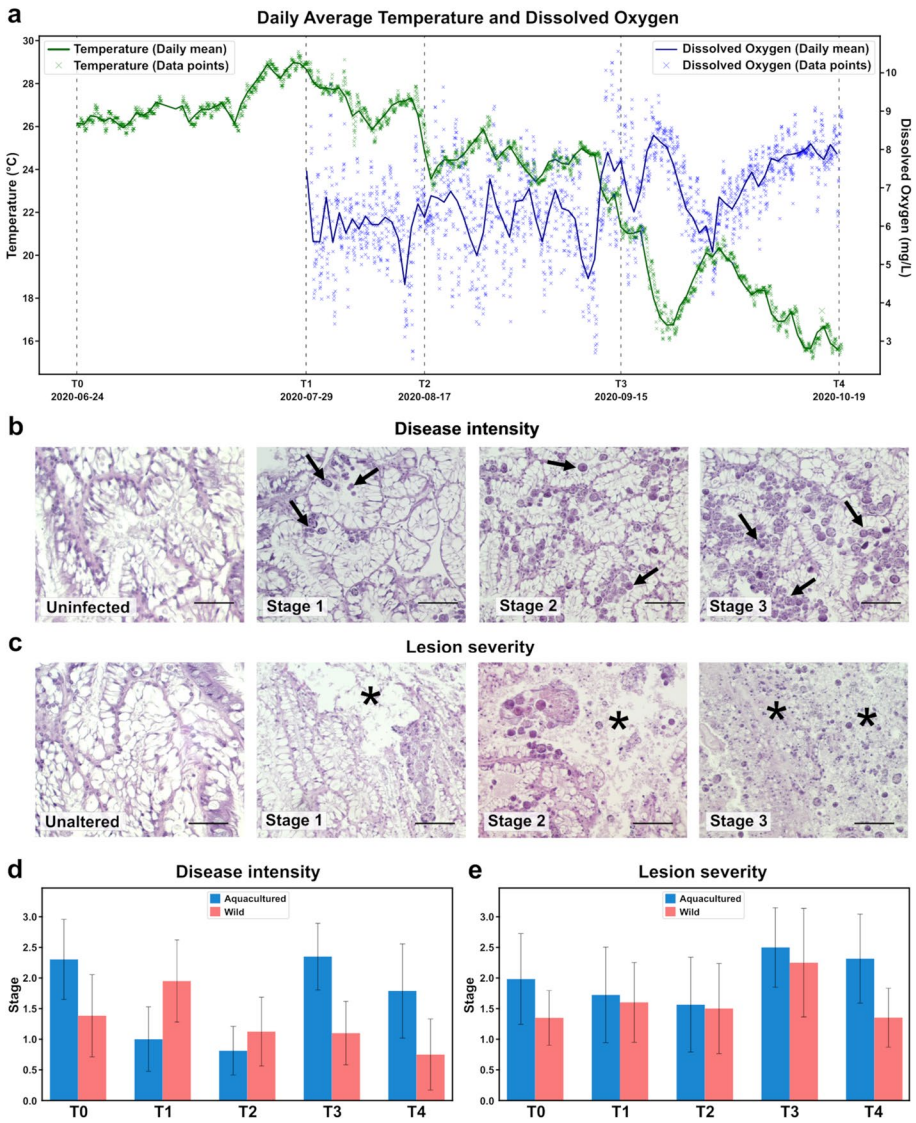

Fig. 1 Environmental and pathological monitoring of experimental scallops. **a** Temperature and dissolved oxygen data from the experimental site. Temperature data before July 29 2020 were collected from a nearby inshore site before the data logger was procured and secured on the scallop bags. **b** Kidney tissue samples from scallops were assessed for disease intensity, which was classified from 0 (uninfected) to 3 (high infection), as indicated by the presence and abundance of parasite cells (denoted by arrows). **c** Severity of lesions within the kidney tissues was evaluated and scored from 0 (unaltered) to 3 (severe alterations), with asterisks marking the tissue lesions. The scale bar applicable to both **b** and **c** is 20 µm. Bar plots summarizing disease intensity **d** and lesion severity **e** in experimental aquacultured and wild scallops, where T0 is the initial time point (June), and T1 to T4 are scallops at 1, 2, 3, and 4 months, respectively. Error bars represent standard deviations

minimal in mid-August with a daytime level of 2.5 mg/L. All scallops (both aquacultured and wild) displayed BSM infections (100% prevalence) at Time 0 (June). Such 100% prevalence of BSM infection in one-year-old scallops at the start of the summer is typical in the Peconic Estuary [25]. Prevalence remained unchanged through the entire experiment for aquacultured scallops but dropped to 75% in October for wild scallops (Supplementary Table S1). Infection intensity (Fig. 1b) showed a bimodal distribution in aquacultured scallops with high intensity (average intensity = 2.30) in June followed by a decrease in July (1.00) and August (0.81), then a marked increase in September (2.35) followed by a slight decline in October (1.79) (Fig. 1d). In contrast, infection intensity in wild scallops was unimodal and increased from 1.38 in June to a maximum of 1.95 in July, before decreasing to 1.13, 1.10 and 0.75 in August, September, and October, respectively (Fig. 1d).

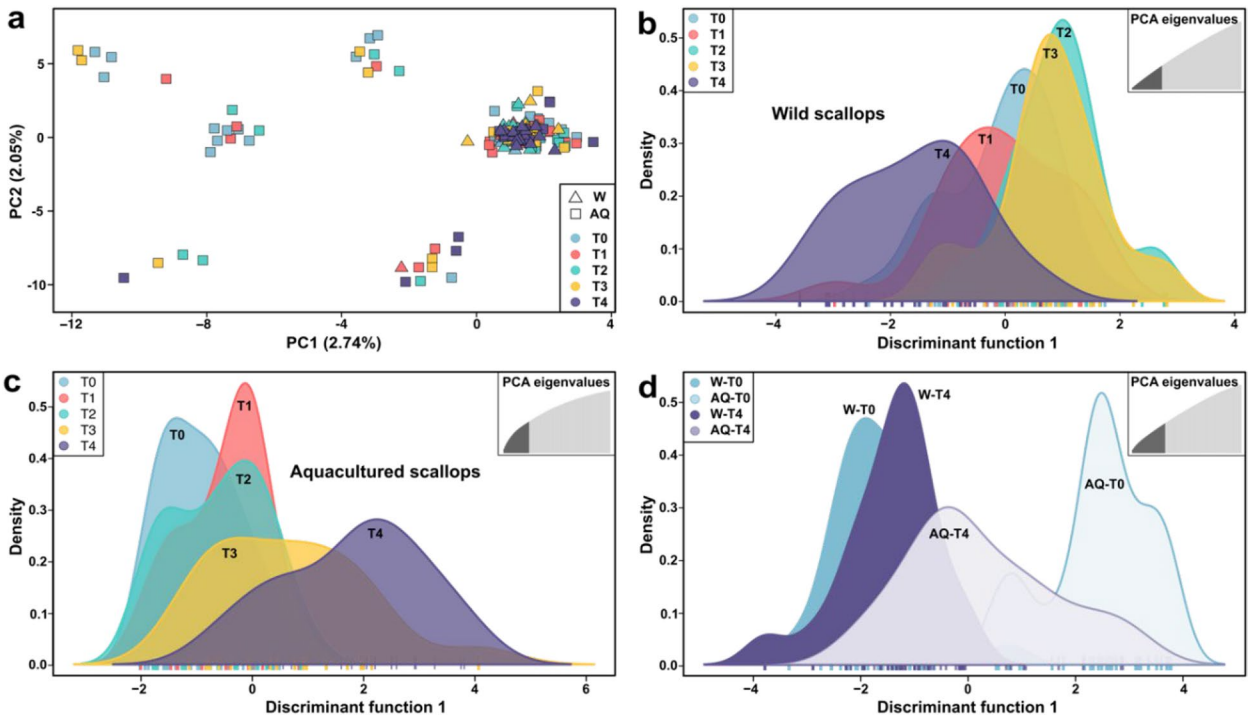

Fig. 2 Comparative analysis of genotypes from wild and aquacultured scallops. **a** Principal Component Analysis (PCA) of all analyzed scallops. Discriminant Analysis of Principal Components (DAPC) of genotypes from wild (**b**) and aquacultured (**c**) scallops where T0 is the initial time point (June), and T1 to T4 are scallops at 1, 2, 3, and 4 months, respectively; **d** both aquacultured and wild scallops at the beginning of the experiment and after four months of cultivation in Flanders Bay

In parallel, kidney lesions (Fig. 1c) were assessed throughout the study and were systematically more severe in aquacultured scallops as compared to wild animals within each of the sampling points. Kidney lesion ranking in aquacultured scallops decreased from 1.98 in June to 1.72 in July and 1.56 in August, before increasing markedly in September (2.50, maximum) and October (2.32). Maximum kidney lesion ranking was also noticed in October among wild scallops (2.25) (Fig. 1e). Deployed scallops suffered significant mortality, particularly among aquacultured stocks where cumulative mortality reached 90.1% in October (Supplementary Table S1). Mortality was significantly lower among the wild stocks (36.3%) deployed under the same experimental conditions. Qualitative diving records showed that mortality in aquacultured scallops was limited through August and peaked between August and September (~60%) and through October suggesting that environmental conditions during summer (Supplementary Table S1) were not the primary driver of bay scallop mortality in this experiment. While significantly lower than mortality levels in aquacultured scallops, mortality among caged wild scallops was also limited in summer but progressed throughout the last sampling event after four months of deployment. Notably, mortality dynamics from dive surveys suggest that the highest levels of mortality occurred when kidney lesions were most severe (September and October).

Scallop genotyping results

In parallel to scallop monitoring throughout the four-month deployment period, we also conducted genotyping using RAD-Seq technology to obtain data on genetic variations. Using RAD-Seq data and the reference genome we picked up 6,667 high-quality SNPs (Supplementary Table S2). These SNPs were used to understand whether the genetic structure of scallops changed under the combined pressure of parasitic infection and environmental stressors throughout the summer mortality event. The principal component analysis (PCA) plot (Fig. 2a) captures the distribution of genetic variation within wild and aquacultured scallops at five consecutive time points (T0 to T4, i.e., June to October) during the mortality event.

The aquacultured scallops display a wider spread in genetic variation, particularly at T0 through T2, indicating a greater genetic heterogeneity among individuals within this group. In contrast, the wild scallops show less dispersion, indicating a more stable genetic makeup over time. The convergence of genetic variation in aquacultured scallops over successive months may reflect the

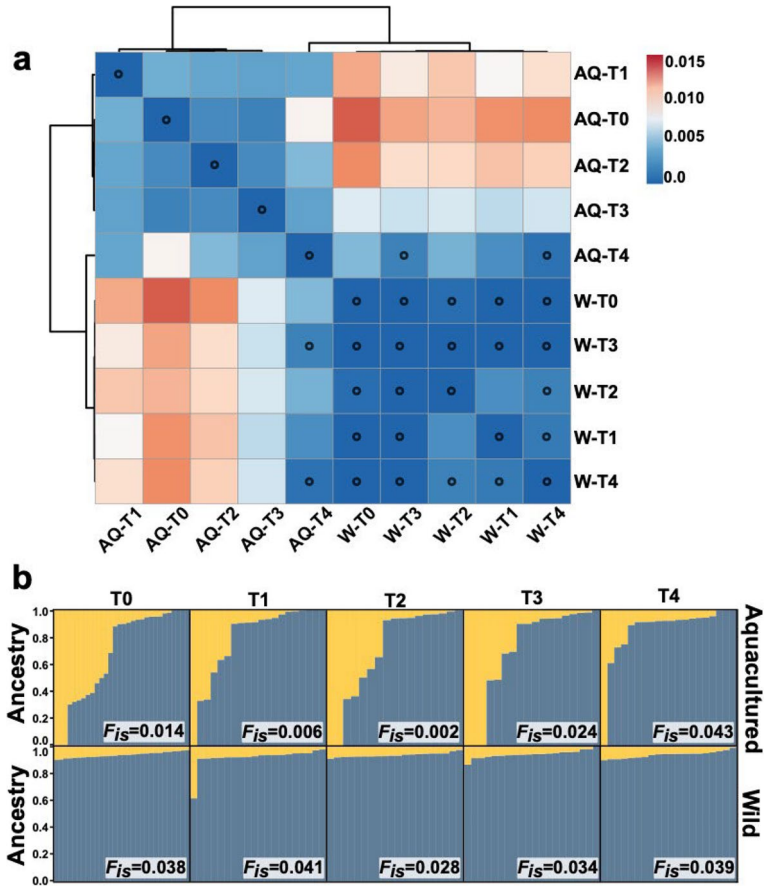

Fig. 3 Genetic variation and population structure in scallops over time. **a** Heatmap of pairwise genetic differentiation F_{ST} values. Each cell represents the F_{ST} value for the intersection of the respective scallop's group on the x and y axes. Values marked with "o" do not show significant differences (p-value > 0.05) or represent the same groups. **b** Population structure (K = 2) and inbreeding coefficient (F_{IS}) analysis. Bar plots detail the population structure levels in wild (W) and aquacultured (AQ) scallops over the study duration. The y-axis quantifies the genome's ancestral proportion. The two-color codes in each bar represent different genetic ancestries. The F_{IS} values are displayed in the lower right corner of each subplot, indicating the level of inbreeding or genetic homogeneity within each group at different time points

impact of selective pressures. Since mortality rates among wild and aquacultured scallops differed, we analyzed them separately. In the case of wild scallops (Fig. 2b), where mortality was low, we noticed only a slight shift in the genetic structure after four months. In the case of the aquacultured scallops, a shift in community structure was much larger and began to occur as early as the third month of observation (Fig. 2c). This is likely due to the high mortality rate of scallops in this group. Interestingly, the genetic make-up of survivor aquacultured scallops at the end of the deployment was closer to that of wild scallops (Fig. 2d). The pairwise F_{ST} analysis conducted over a four-month period confirmed these trends and revealed notable changes in the genetic differentiation between aquacultured and wild scallops throughout the mortality event (Fig. 3a).

The heatmap visualization of F_{ST} values across sequential time points demonstrates a clear trend towards genetic homogenization between the two cohorts. Initially, the aquacultured scallops (AQ) presented a distinct genetic profile as compared to the wild (W), with higher F_{ST} values indicating substantial genetic differentiation. As the study progressed, these values consistently decreased, particularly after the second month (T2), suggesting a shift in the genetic structure of the aquacultured group. By the fourth month (T4), the F_{ST} values between T4_AQ and T4_W scallops suggested a marked reduction in genetic differentiation, indicating that the surviving aquacultured scallops had become genetically more similar to the wild. The heatmap also reveals intra-population genetic stability in the wild scallops, with relatively low F_{ST} values throughout the study duration, further emphasizing the genetic resilience of the wild cohort. In contrast, the aquacultured scallops shows a dynamic shift, ending with a genetic composition that mirrors

the wild animals more closely than at the outset of the experiment.

To identify genetic groups that may have varying resistance to parasitic infection, we conducted Structure analysis, utilizing K = 2 (Fig. 3b) as determined by the MedMeak, MaxMeak, MedMedK and MaxMedK metrics (Table S3). Aquacultured scallops, initially with a higher proportion of certain genetic ancestry, showed a significant decline in this component (colored yellow), correlating with increased mortality. This pattern indicates a selective disadvantage of these genetic traits when challenged by BSM infection in the environmental conditions of Flanders Bay. In contrast, the wild scallops maintained a consistent genetic admixture profile, suggesting a stable genetic structure that could be indicative of greater resilience. By the end of the study, the genetic composition of aquacultured survivors more closely resembled that of the wild scallops, implying a selection favoring genetic characteristics associated with higher resistance to the deleterious effects of BSM infection.

The initial groups had positive F_{IS} values of 0.014 and 0.038, respectively. This indicates the presence of inbreeding within both the aquacultured and wild scallops prior to the disruption caused by BSM. The wild group had a higher F_{IS} score, indicating a higher degree of inbreeding compared to the aquacultured. After the BSM outbreak, the surviving groups also had positive F_{IS} values (0.043 and 0.039, respectively) after 4 months.

To better understand the genetic diversity within the aquacultured and wild scallop populations, the observed heterozygosity (*Ho*) and expected heterozygosity (*He*) were analyzed (Table S4). The results indicate that heterozygosity levels within each group were relatively stable over time. At T0, aquacultured scallops exhibited significantly higher *Ho* (0.212 ± 0.0019) compared to wild scallops (0.208 ± 0.0018; $p = 0.008$), suggesting greater initial genetic heterogeneity within the aquacultured group. In contrast, *He* values were not significantly different (0.213 ± 0.0017 vs. 0.215 ± 0.0017; $p = 0.159$). Over the following months, *Ho* remained significantly different at T1 ($p = 4.76 \times 10^{-11}$), T2 ($p = 0.002$), and T3 ($p = 2.12 \times 10^{-7}$), but the values converged at T4 (0.2084 ± 0.0019 in aquacultured vs. 0.2069 ± 0.0019 in wild; $p = 0.317$). In contrast, *He* values remained stable at T1 and T2 before diverging at T3 ($p = 1.94 \times 10^{-7}$) and T4 ($p = 0.019$). These findings are consistent with the observed shift in genetic structure over the study period, where initially more diverse aquacultured scallops experienced selection pressures that reduced genetic heterogeneity, leading to a final genetic composition more similar to the wild population.

To investigate genetic relationships within and between the aquacultured and wild scallop populations, we analyzed patterns of shared ancestry using haplotype data with fineRADstructure (Fig. 4). The analysis identified distinct clusters of individuals with elevated levels of recent shared ancestry, predominantly within the aquacultured group. In contrast, the wild scallops had a more uniform genetic distribution with fewer distinct clusters, reflecting a more homogeneous genetic structure. Based on the co-ancestry matrix, six distinct genetic groups were identified, most of which were restricted to the aquacultured group. A small number of individuals showed partial overlap with the wild scallops, suggesting a common ancestry between these groups. Among the aquacultured scallops, the identified genetic groups showed varying degrees of differentiation from each other, reflecting the greater genetic heterogeneity within these animals.

Following the identification of six distinct genetic groups, inter-individual relatedness (Fig. 5a) illustrates that some of them (genetic groups 1, 4, 5, and 6) are characterized by higher internal relatedness, suggesting closer genetic relationships within these groups. This variation underscores the heterogeneity of genetic relationships within the aquacultured cohort.

Temporal dynamics of these genetic groups are presented in Fig. 5b, showing changes in their proportions across the study period from T0 to T4. At T0, certain genetic groups, such as group 6, represented 30% of aquacultured individuals, while group 5 accounted for 10%. By T4, these genetic groups were no longer present, suggesting they were more susceptible to the mortality outbreak. The disappearance of these genetic groups indicates that their genetic traits may yield reduced survival under the prevailing selective pressures. However, it cannot be ruled out that non-genetic factors, such as maternal provisioning or early developmental differences may have also influenced survival outcomes among aquacultured scallops. To further explore the genetic basis of this differential resilience, OutFLANK, pcadapt, and GWDS analyses were performed to identify outlier SNPs associated with the less resilient genetic groups. pcadapt allowed the identification of 53 outlier SNPs, 5 of which were also identified by OutFLANK including 2 SNPs also identified by GWDS (Supplementary Fig. S1, Table S5). Of the five SNPs identified by OutFLANK (and pcadapt), two were intergenic, while two were located in introns—one in a gene encoding an uncharacterized protein and another in the ribitol-5-phosphate xylosyltransferase 1-like gene. The fifth SNP, which was also identified by GWDS, represented an exonic nonsynonymous variant in the nephrocystin-3-like gene.

Allelic frequency changes in scallop groups

In this study investigating the impact of BSM infection on scallop population, we noted significant allele frequency

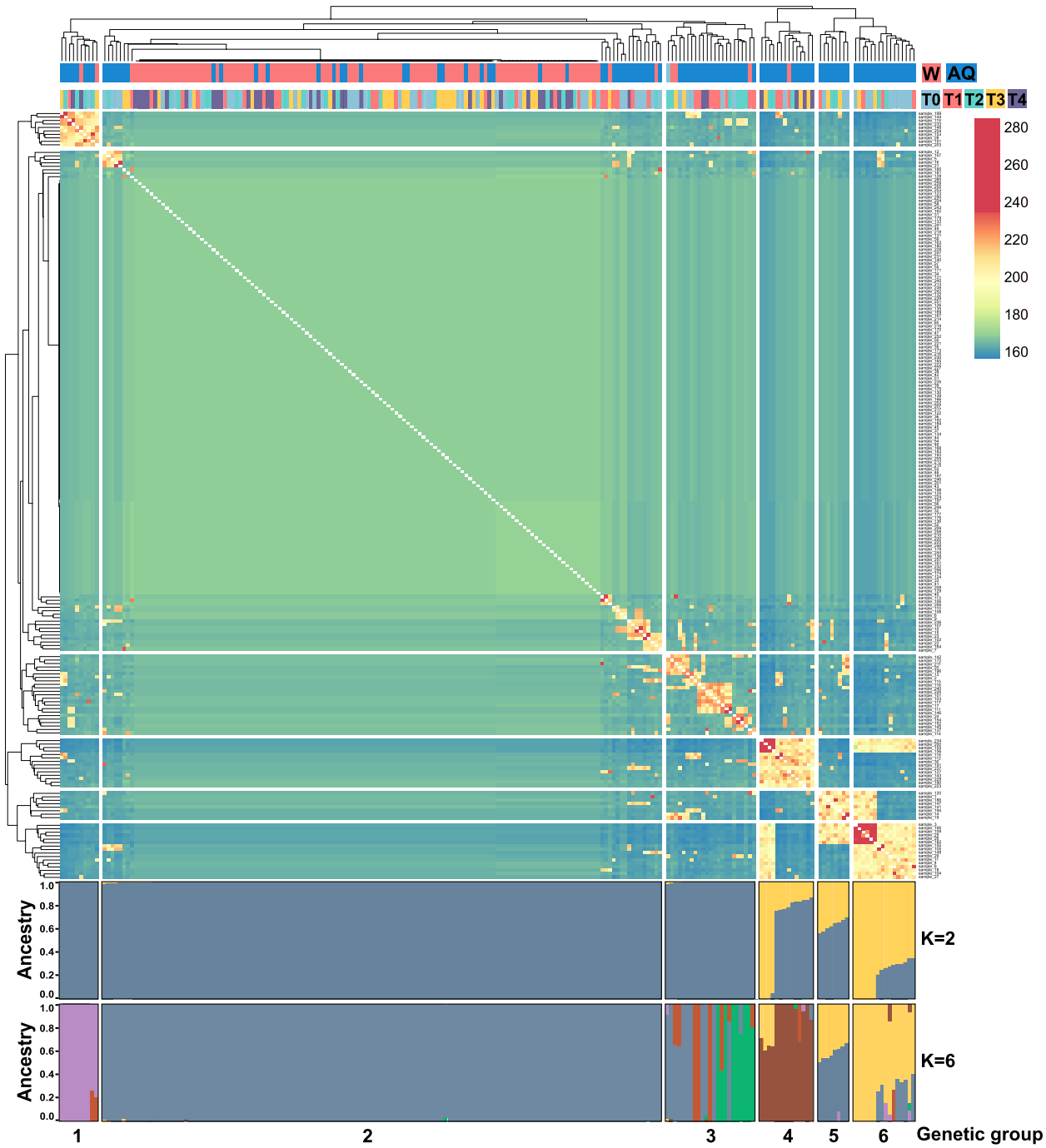

Fig. 4 Genetic similarity and population structure in wild and farmed scallops. The heatmap shows the similarity between pairs of individuals based on shared haplotypes, with colors indicating the proportion of shared haplotypes (red: highest similarity, blue: lowest). Dendrograms show the clustering of individuals based on genetic similarity. Annotations for each sample include the time point (T0-T4) and its designation as wild (W) or aquacultured (AQ). The lower panels show the results of the fastSTRUCTURE analysis for K = 2 and K = 6, illustrating the ancestry proportions of individuals. Each bar represents an individual, with colors indicating genetic groups assignments

shifts in multiple genes involved in stress and infection response. Comparative analysis of initial and surviving scallops over a four-month period revealed substantial changes in SNP allele frequencies (Supplementary Table S6). While significant allele frequency shifts were observed using Fisher's exact test with corrections for multiple comparisons,

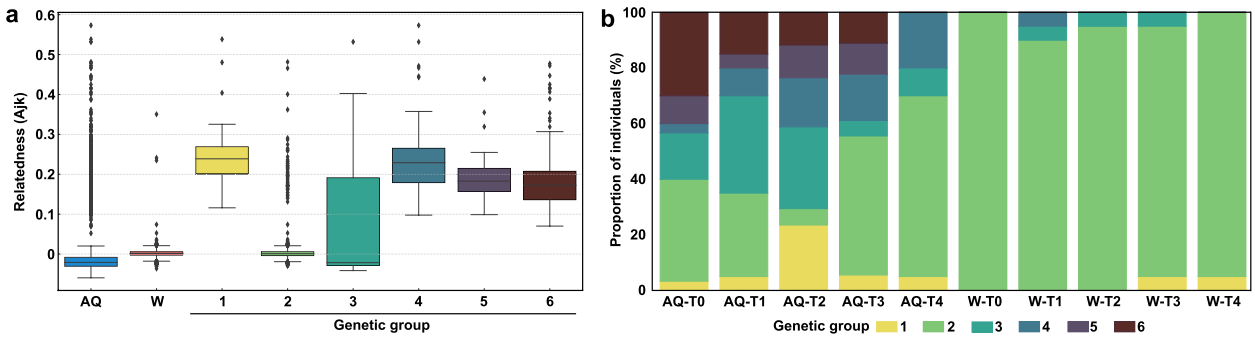

Fig. 5 Inter-individual relatedness and genetic group dynamics in scallops. **a** Boxplot displaying inter-individual relatedness values for all pairs of individuals within each genetic group, calculated using the Ajk method across different time points. **b** Stacked bar plot showing the percentage of individuals belonging to different genetic groups within the wild (W) and aquacultured (AQ) groups at each time point

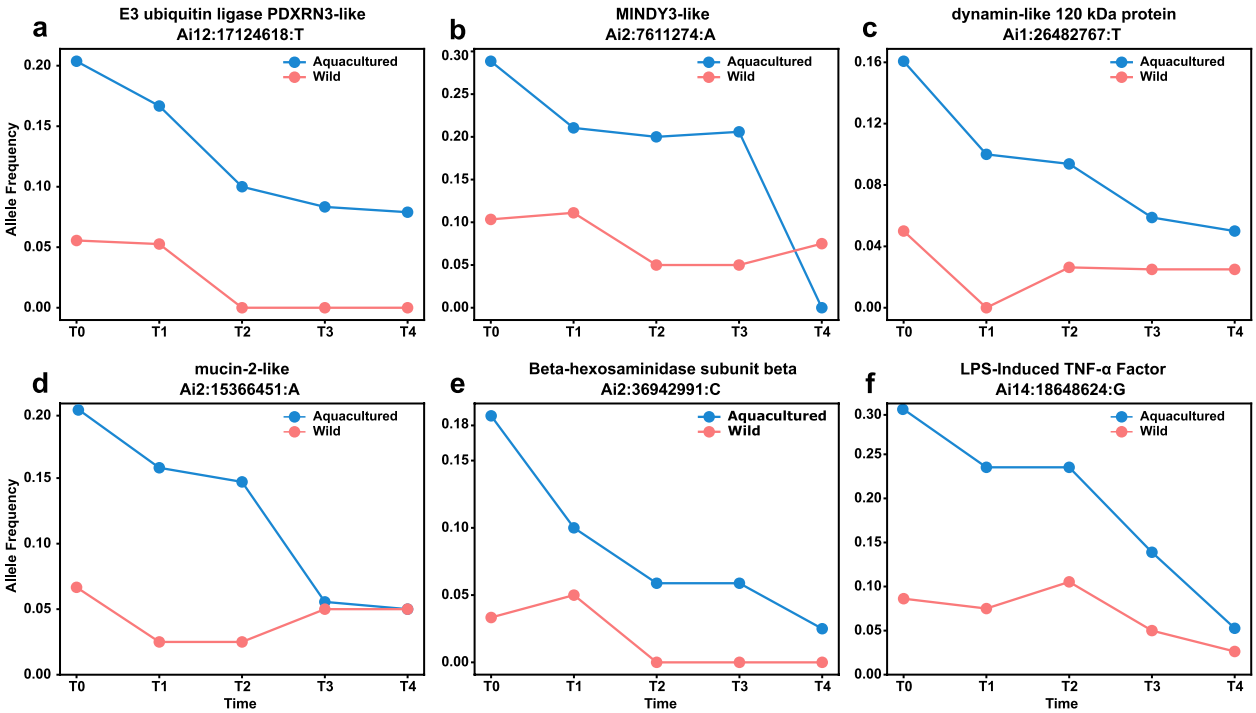

Fig. 6 Temporal allele frequency changes in scallop groups. **a-f** plots represent allele frequency trajectories in genes related to stress and immune response in wild and aquacultured scallops over a 4-month period. Notable allelic shifts are observed in genes encoding protein ubiquitination, metabolism, and immune response. The plots compare frequencies at baseline (T0) and monthly intervals (T1-T4)

only one SNP passed the stringent Bonferroni threshold ($p < 0.05$), and 13 SNPs met the criteria under the Benjamini–Hochberg correction (FDR < 0.05). Several of these SNPs are located within genes involved in protein ubiquitination, metabolism, and immune responses. Notable examples include genes coding for E3 ubiquitin-protein ligase PDZRN3-like, ubiquitin carboxyl-terminal hydrolase MINDY-3-like, beta-hexosaminidase, lipopolysaccharide-induced tumor necrosis factor-alpha homolog and dynamin-like 120 kDa protein (Fig. 6).

Particularly in the gene encoding the E3 ubiquitin ligase PDZRN3-like a cytosine to thymine (C > T) variant was detected (Fig. 6a). Over the 4-month period, the allele frequency in wild and aquacultured scallops showed a marked decrease. The allelic variant "T", which was initially more prevalent in the aquacultured group at time zero (T0), exhibited a steady decline in frequency at each subsequent time point (T1-T4). A similar allele frequency reduction was observed in the *mindy3-like* gene, with an adenine to guanine (A > G) shift (Fig. 6b), and

in the dynamin-like 120 kDa protein gene, which exhibited a C > T transition (Fig. 6c). This trend of decreasing allele frequencies in aquacultured scallops was further evidenced in the mucin-2-like gene, where a guanine to adenine (G > A) transition was observed (Fig. 6d). In the beta-hexosaminidase subunit beta gene, the G > C polymorphism became undetectable in wild scallops by T2, with a stark reduction from 0.18 at T0 to 0.03 at T4 for aquacultured animals (Fig. 6e). The LPS-induced TNF-alpha factor homolog gene exhibited an A > G change, with the allele frequency decreasing from 0.31 to 0.06 in aquacultured scallops, over the study duration (Fig. 6f).

Discussion

Building on the genetic insights that suggest an adaptive genomic response in *A. irradians*, our findings from the field study provide a real-world context where these adaptations may manifest. The differential responses to BSM infection and environmental stress between wild and aquacultured scallops, as evidenced by variations in infection intensities, kidney lesion severities, and mortality rates, are suggestive of underlying genetic factors that confer resilience. The genetic diversity and heterogeneity observed in aquacultured scallops highlights the complex interplay between cultivating practices and restoration efforts in the Peconic Estuary. In this context, the unexpected pattern of higher heterogeneity in cultured populations compared to wild populations at T0 can be explained by the historical reliance on hatchery-produced scallops for restoration, resulting in genetic overlap between the two groups. This finding highlights the challenge of maintaining clear distinctions between wild and aquacultured groups, particularly in regions where aquaculture has been integral to species recovery. In addition, the possible genetic bottlenecks experienced by wild scallops over the last two decades as a result of high mortality rates related to harmful algae blooms may have contributed to their lower observed diversity, as suggested by lower observed heterozygosity. Moreover, the high mutation rates often observed in aquacultured group could also influence genetic patterns and should be considered as a contributing factor [27]. Therefore, the observed clustering of aquacultured individuals, as revealed by the PCA and relatedness analyses, may be attributed to founder effects and family structure commonly reported in hatchery-reared stocks. Similar patterns have been observed in other bivalve species, such as Yesso scallops, where aquaculture practices led to reduced genetic diversity and increased differentiation between wild and cultured populations due to low effective population sizes and limited broodstock variability [28]. The presence of such structured clustering within the aquacultured scallops emphasizes the influence of aquaculture practices on genetic diversity, potentially leading to higher susceptibility to environmental stresses, as indicated by the selective pressures from BSM infection observed in this study. The convergence of F_{ST} values over time between the aquacultured and wild scallops likely reflects the selective mortality pressure exerted by the BSM infection possibly in conjunction with stressful field conditions, which has disproportionately culled the genetically susceptible individuals within the aquacultured group. The alignment of genetic profiles suggests that the traits contributing to survival in the wild scallops, presumably encompassing resistance to the combined stress of BSM and summer environmental conditions, are being selected for within the aquacultured animals. The F_{ST} values in the aquacultured scallops show a consistent pattern over time, indicating a bottleneck effect that reduces genetic diversity and makes the population more homogeneous. Notably, the F_{IS} values increased from the baseline population, suggesting that the surviving individuals in both groups may have experienced an increase in homozygosity (inbreeding) due to the selective pressure exerted by the BSM outbreak under prevailing environmental conditions. Survivors may possess specific genetic traits that allowed them to resist the BSM infection, leading to a decrease in genetic diversity and an increase in homozygosity within groups. This perspective assumes that the BSM outbreak played a significant role in shaping the genetic structure of the scallop population.

Interestingly, our results showed that the three outlier SNP analysis tools used (pcadapt, OutFLANK, and GWDS) converged and allowed the identification of two SNPs (identified using all three methods), one of which representing an exonic nonsynonymous variant in the nephrocystin-3-like gene. This gene stands out for its recognized role in cilia activity in epithelia, kidney function and renal diseases in model species [29, 30], including fish [31]. Alterations of nephrocystin genes have also been associated with abnormal excretory (i.e., renal) functions in *Caenorhabditis elegans*, suggesting mechanistic conservation across the Metazoa [32]. These observations strongly support the involvement of this gene in scallop resilience/susceptibility during BSM infection of scallop kidneys although additional studies are needed to confirm this mutation as a causative factor for higher survivorship/mortality. Regarding the genes in which significant changes in SNP allele frequency have been observed, they also do not appear to be random due to their involvement in resistance to infection and stress. The E3 ubiquitin-protein ligase *pdzrn3-like* gene, could play a role in the immune response or stress adaptation given the gene's involvement in protein ubiquitination, a process crucial for protein turnover and signaling [33, 34]. Mutations in *mindy3-like* gene could reflect

differential survival strategies related to these genes, which are integral to protein degradation and regulatory pathways in cellular stress responses [35]. The dynamin-like 120 kDa protein gene is essential for vesicle trafficking and membrane remodeling, processes critical for cellular responses to stressors [36, 37]. The mucin-2-like gene plays a crucial role in mucosal immunity and barrier functions [38, 39]. The beta-hexosaminidase subunit beta gene is essential for lysosomal degradation pathways and the LPS-induced TNF-alpha homolog gene is involved in the innate immune response [40]. All these genes are important in regulating the inflammatory pathways and may contribute to scallop survival under pathogenic stress. But it should be mentioned that SNPs identified by RAD-Seq in these genes are located within intronic regions or as non-synonymous exonic variants of their respective genes. Furthermore, it's important to note that RAD-Seq may not capture the causative SNPs underlying specific traits, but rather SNPs in linkage with the causative SNP(s) [41]. The SNPs, therefore, represent putative markers genetically linked to the traits of interest, warranting further investigation to pinpoint the exact mutations and their effects on scallop physiology and survival. Similarly, SNPs displaying significant changes in frequency throughout the mortality event include many intergenic features, but also encompass uncharacterized genes, and further investigations are needed to unravel how these variations may enhance scallop survivorship. In this framework, our results provide an exploratory framework for understanding genetic shifts under the combined pressures of BSM infection and environmental stress. While allele frequency changes were statistically significant for some SNPs, it is important to acknowledge that these results are preliminary and reflect trends that require further validation through additional studies. Overall, this study is the first to explore genetic adaptation in bay scallops in relation to resilience to BSM infection and represents an important step towards identifying potential genomic markers associated with this resilience.

Nevertheless, results presented here suggest a complex genetic adaptation process in scallops, driven by the selective pressures of BSM infection. The wild scallops, displaying more stable allele frequencies, may possess a genetic makeup that confers resilience to environmental and pathogenic stresses translated into a much higher survivorship than aquacultured scallops. These alterations, particularly in genes associated with immune response and stress tolerance, may indicate a process of natural selection that favors individuals with certain genetic traits, although additional research is needed to firmly establish these conclusions. But similar alterations in genetic structure have been reported in other marine organisms. For example, studies of coral populations, such as those from the Great Barrier Reef, have documented significant changes in genetic composition due to bleaching events associated with climate change [42]. In addition, Atlantic cod populations along the Canadian coast have shown genetic shifts related to overfishing and subsequent environmental stress [43]. In the case of terrestrial species, the European ash tree has shown rapid changes in allelic frequencies in response to ash dieback, a serious problem caused by the fungal pathogen *Hymenoscyphus fraxineus* [44]. Our own previous research demonstrated changes in SNP frequencies in the hard clam (*Mercenaria mercenaria*) following mortalities induced by a microbial infection [45]. These examples highlight a broader biological context in which genetic diversity plays a key role in the survival and adaptability of species under environmental or disease pressure.

While our results allowed the identification of genetic features putatively involved in scallop survivorship under BSM infection pressure, we recognize that the lack of replicated populations limits the generalizability of our conclusions. Our study focused on a single site, selected for its high disease prevalence, environmental stress and high mortality levels, providing a critical context for examining genetic responses under severe selective pressures. However, conducting experiments across a wider range of locations with varied mortality levels may enhance our understanding of the relationships between the genetic background of scallops and their survivorship. Future studies should aim to include multiple populations from different geographic areas to evaluate whether the trends noted here hold universally. Such studies would provide a more comprehensive framework for understanding the genetic determinants underlying resilience in scallop populations.

Conclusions

The economic and ecological importance of bay scallops requires a deeper understanding of their genomic structure and resilience. Insights from this study contribute to understanding genetic resilience in scallops, which is an important step toward informing strategies for managing populations, supporting sustainable aquaculture practices, and enhancing breeding programs. This is particularly relevant in the context of escalating environmental change and its impact on marine biodiversity. This study explores genetic resilience mechanisms and host–pathogen interactions in scallops, providing a basis for further research. By combining genomic findings with environmental and epidemiological analyses, the study also highlights its potential to inform breeding strategies in aquaculture, aimed at improving scallop stock resilience and productivity. A non-synonymous mutation in a gene

involved in kidney functions in vertebrates and invertebrates has been associated with higher scallop survivorship and may represent a prime candidate for future targeted investigations. Overall, our findings underscore the importance of genomic science in managing marine resources and implementing conservation strategies. This approach lays the groundwork for future studies into the genetic basis of adaptation and resilience in bay scallops, emphasizing the value of genomics in mitigating the challenges posed by environmental changes and pathogen pressures.

Materials and methods

Sample collection

Aquacultured scallops (54.5 ± 3.8 mm in shell width, mean ± standard deviation) were sourced from a stock farmed by Cornell Cooperative Extension in Orient Harbor, New York, USA (41.137904, −72.315392). Wild scallops (56.7 ± 4.6 mm) were collected from the same site and therefore were exposed to the same oceanographic conditions. Both stocks were separately deployed in Flanders Bay (40.917634, −72.593486), a site characterized by the presence of BSM and where mortality pressure was high during the previous summer (estimated to exceed 90% among adult scallops). Each stock was deployed in four replicate plastic mesh bags (91.5 cm × 40.5 cm × 11.5 cm rectangular cages made of 1-cm square mesh; 62 scallops per bag). A data logger (miniDOT, Precision Measurement Engineering, Inc., Vista, California, USA) was attached to one of the replicate bags to collect temperature and dissolved oxygen data (for logistical reasons the logger was deployed 1 month after scallop deployment). A comprehensive analysis was conducted on a total of 220 wild and aquacultured scallops using RAD-Seq. The study involved collecting scallops at Time 0, followed by monthly collections over a four-month period post-deployment. RAD-Seq was performed on 30 scallops per cohort at Time 0 and 20 scallops from each cohort at subsequent time points to genetically characterize scallops present in NY and evaluate if field mortality exerted a selective pressure based on scallop genetic makeup. All sampled scallops were submitted to standard histopathological processing as described below.

Histopathology

All scallops used for RAD-Seq were also submitted to standard histopathological processing to evaluate BSM infection and tissue conditions, Briefly, representative sections of scallop tissues including kidney, gonad, adductor muscle, digestive tract, gills, and mantle were fixed, embedded in paraffin, sectioned and stained with hematoxylin and eosin following standard protocols. Microscopic observations of tissues evaluated the presence of BSM and lesions in kidney tissue as described previously [25]. Disease intensity was ranked from 0 (absence of BSM) to 3 (abundant BSM cells in kidney tissues) on a semi-quantitative scale. Similarly, kidney lesions were ranked from 0 (intact kidney tubule structure) to 3 (severe lesions associated with disruption of kidney epithelium and disorganization of tubule structure).

RAD-Seq library preparation and sequencing

DNA was extracted from the adductor muscle of a total of 220 individual scallops (see "Sample collection" section) using a standard phenol:chloroform:isoamyl extraction [46]. DNA quality and quantity were checked using the Qubit dsDNA HS Assay Kit with assays read on a Qubit 2.0 fluorometer (Life Technologies/Agilent 2100; Waldbronn, Germany). DNA was then submitted to digestion using 3 enzyme combinations (PstI/MseI, KasI/AciI and KasI/HpyCH4IV) following a protocol with steps modified from Peterson et al. [47]. Following digestion, DNA samples were individually ligated to specific P1-n barcoded adapters, size selected using Nucleomag NGS clean-up and size select kit (Macherey–Nagel, Germany). Adapter-ligated fragments were then PCR-amplified with uniquely indexed PCR-P2 primer sequences to allow for post-sequencing demultiplexing. PCR products were then quantified using Qubit dsDNA HS assay kit, and then size-selected with Pippin-Prep (1.5% agarose dye free gel cassette, Sage Science, Beverly, Massachusetts, USA). Samples were sequenced using the Illumina NovaSeq 6000 System with an S4 flow cell. A total of 217 samples were successfully sequenced (Supplementary Table S7). Demultiplexed reads were processed using Stacks v. 2.59 [48] as described in the protocol of Rochette and Catchen [49]. Reads were checked for correct cut sites and adapter sequences using the 'process_radtags' module, and then mapped to the chromosome-level genome assembly of the bay scallop *Argopecten irradians* NY (GCA_041381155.1) [50] using minimap2 v. 2.17 [51]. Sorted and indexed mapping reads were generated using SAMtools v. 1.17 [52]. The use of the 'gstacks' module [48] facilitated the SNP calling process, and subsequently the 'populations' module enabled the export of these called variants in VCF format. Following the guidelines of O'Leary et al. [53], the VCF file containing the genotypes was filtered using VCFtools v. 0.1.16 [54], BCFtools v. 1.17 [52], and vcfR v. 1.14.0 [55]. The SNP data filtering scripts have been adapted from those available at https://github.com/chollenbeck/king_scallop_popgen_2022 [56] using the same filtering parameters. Briefly, SNPs with depth < 10 and genotype quality < 20 were excluded on a per-genotype basis. Sites with a minor allele frequency below 0.05 or more than 10% missing data in any group were removed. Finally, SNPs deviating from Hardy–Weinberg equilibrium ($p < 0.001$) were filtered,

and thinning was applied to reduce linkage disequilibrium (1000 bp window). Highly related individuals were identified using the relatedness calculation function (–relatedness) as implemented in VCFtools. A threshold of Ajk > 0.7 was applied to detect pairs of individuals with abnormally high relatedness [57], and two samples (Table S8) were removed to reduce the impacts of family structure on PCA [58–60]. Beagle v. 5.4 [61] was used for the phasing of SNP genotypes. Linkage disequilibrium (LD) pruning of the phased VCF file was performed using the snp_autoSVD function in bigsnpR v. 1.12.2 [62]. The diveRsity R package v. 1.9.90 was used to calculate observed heterozygosity (*Ho*), expected heterozygosity (*He*), and inbreeding coefficient within individuals (F_{IS}) for each dataset using only on polymorphic loci. The statistical significance of differences in *Ho* and *He* values between aquacultured and wild scallop populations at different time points was determined using t-tests. PCA was performed using PLINK v. 1.90 (scree plot presented in Supplementary Fig. S2) [63]. Population differentiation was assessed by calculating pairwise FF_{ST} values using the 'stamppFst' function in the R package StAMPP v. 1.6.3 [64]. Structure analysis was performed with fastSTRUCTURE v. 1.0 [65] using a logistic prior and cross-validation (cv = 10) over values of K from 1 to 10 to estimate the most likely number of genetic clusters. The optimal number of clusters (K) was determined using the MedMeaK, MaxMeaK, MedMedK and MaxMedK metrics [66] via the StructureSelector web tool (https://lmme.ac.cn/StructureSelector) [67]. Cluster partitions were then visualized using pophelper [68]. Population structure was also inferred by calculating the similarity between each individual's haplotype using the fineRADStructure package [69] to estimate co-ancestry between individuals. Outlier detection was performed using OutFLANK [70], pcadapt [71], and GWDS [72] implemented in the SambaR package v. 1.10 [73]. The visualization of the population structures was performed by a Discriminant Analysis of Principal Components (DAPC) using adegenet v. 2.1.10 [74]. Cross-validation determined the number of principal components (PCs) retained in each analysis. Fisher's exact test [75] was performed to evaluate the significance of changes in allele frequencies between T0 and T4, with multiple comparisons accounted for using the Benjamini-Hochberg (BH) method [76] to control the false discovery rate and the Bonferroni correction [77] as a more conservative measure.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-11337-y.

Supplementary Material 1.

Supplementary Material 2.

Acknowledgements
Authors are very thankful to the members of the Marine Animal Disease Laboratory (Madison Muehl, Younes Bouallegui, Sabrina Geraci-Yee, Caroline Schwaner, Ray Czaja, Christopher Brianik and others) for help with scallop processing and to Cornell Cooperative Extension team members (Scott Hughes, Kaitlyn O'Toole and others) for field support.

Authors' contributions
E.P.E, B.A. and S.T. designed the study and secured the funding. S.T. and H.T collected biological samples. A.T. and I.B. generated RAD-Seq libraries. D.G. performed genomic analysis. D.G., E.P.E. and B.A. analyzed data and drafted the paper. All authors contributed to the editing of the manuscript and approved the final version of the paper.

Funding
This work was primarily supported by NSF Grant number IOS-2026358 to B.A., E.P.E. and S.T. Financial support was also provided by the McConnell Family Foundation and by the New York State Department of Environmental Conservation.

Data availability
All data generated or analyzed during this study are included in this published article and its supplementary files or shared on NCBI under BioProject PRJNA1050236.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY 11794-5000, USA. [2]Cornell Cooperative Extension of Suffolk County, Southold, NY 11971, USA. [3]Station Biologique de Roscoff, CNRS/Sorbonne Université, Place Georges Teissier, 29680 Roscoff, France.

Received: 31 May 2024 Accepted: 10 February 2025
Published online: 15 February 2025

References

1. Byers JE. Effects of climate change on parasites and disease in estuarine and nearshore environments. PLoS Biol. 2020;18: e3000743.
2. Hoegh-Guldberg O, Bruno JF. The impact of climate change on the world's marine ecosystems. Science. 2010;328:1523–8.
3. Burge CA, Kim CJS, Lyles JM, Harvell CD. Special issue oceans and humans health: the ecology of marine opportunists. Microb Ecol. 2013;65:869–79.
4. Ruppert EE, Fox RS, Barnes RD. Invertebrate zoology: a functional evolutionary approach. 7th ed. Belmont, CA: Thomson-Brooks/Cole; 2004.
5. Pechenik JA. Biology of the invertebrates. 5th ed. Boston: McGraw-Hill, Higher Education; 2005.
6. Barnes RSK, Calow P, Olive PJW. The invertebrates: a new synthesis. 2nd ed. Oxford [England] ; Boston: Blackwell Scientific; 1993.
7. FAO, editor. The State of World Fisheries and Aquaculture 2022. Towards Blue Transformation. Rome: FAO; 2022.
8. Sousa R, Gutiérrez JL, Aldridge DC. Non-indigenous invasive bivalves as ecosystem engineers. Biol Invasions. 2009;11:2367–85.
9. Bódis E, Tóth B, Sousa R. Massive mortality of invasive bivalves as a potential resource subsidy for the adjacent terrestrial food web. Hydrobiologia. 2014;735:253–62.
10. Bert TM, Arnold WS, McMillen-Jackson AL, Wilbur AE, Crawford C. Natural and anthropogenic forces shape the population genetics and recent evolutionary history of eastern United States bay scallops (*Argopecten irradians*). J Shellfish Res. 2011;30:583–608.

11. Qin Y, Liu X, Zhang H, Zhang G, Guo X. Identification and mapping of amplified fragment length polymorphism markers linked to shell color in bay scallop, *Argopecten irradians irradians* (Lamarck, 1819). Mar Biotechnol. 2007;9:66–73.
12. Estabrooks SL. The possible role of telomeres in the short life span of the bay scallop, *Argopecten irradians irradians* (Lamarck 1819). J Shellfish Res. 2007;26:307–13.
13. Fusui Z, Yichao H, Xiangsheng L, Jianghu M, Shuying L, Lingxin Q. Introduction, spat-rearing and experimental culture of bay scallop, *Argopecten irradians* Lamarck. Chin J Ocean Limnol. 1991;9:123–31.
14. Yu L, Tan X, Guan X, Mu Y, Lam V, Sumaila R. Value chain of the data-poor Chinese bay scallop aquaculture. Mar Policy. 2023;150: 105556.
15. Guo X, Luo Y. Scallops and scallop aquaculture in China. In: Developments in Aquaculture and Fisheries Science. Elsevier; 2016. p. 937–52.
16. Belding DL. A report upon the scallop fishery of Massachusetts: including the habits, life history of *Pecten Irradians*, its rate of growth, and other facts of economic value. Wright & Potter; 1910.
17. Bricelj MV, Epp J, Malouf RE. Comparative physiology of young and old cohorts of bay scallop *Argopecten irradians irradians* (Lamarck): mortality, growth, and oxygen consumption. J Exp Mar Biol Ecol. 1987;112:73–91.
18. Tettelbach. Seasonal changes in a population of northern bay scallops, *Argopecten irradians irradians* (Lamarck, 1819). In: An international compendium of scallop biology and culture. Baton Rouge, Louisiana: World Aquaculture Society; 1991. p. 164–75.
19. Carroll, John Michael. The effects of habitat and predation on bay scallop populations in New York. Doctoral Dissertation. State University of New York at Stony Brook; 2012.
20. Tettelbach S, Peterson B, Carroll J, Furman B, Hughes S, Havelin J, et al. Aspiring to an altered stable state: rebuilding of bay scallop populations and fisheries following intensive restoration. Mar Ecol Prog Ser. 2015;529:121–36.
21. Tettelbach S, Czaja R, Tobi H, Hughes S, Peterson B, Heck S, et al. Collapse of the New York Bay scallop fishery despite sustained larval and juvenile recruitment. Mar Ecol Prog Ser. 2023;714:45–56.
22. Cosper EM, Dennison WC, Carpenter EJ, Bricelj VM, Mitchell JG, Kuenstner SH, et al. Recurrent and persistent brown tide blooms perturb coastal marine ecosystem. Estuaries. 1987;10:284–90.
23. Tettelbach ST, Wenczel P. Reseeding efforts and the status of bay scallop Argopecten irradians (Lamarck, 1819) populations in New York following the occurrence of" brown tide" algal blooms. J Shellfish Res. 1993;12:423–31.
24. Tettelbach ST, Peterson BJ, Carroll JM, Hughes SW, Bonal DM, Weinstock AJ, et al. Priming the larval pump: resurgence of bay scallop recruitment following initiation of intensive restoration efforts. Mar Ecol Prog Ser. 2013;478:153–72.
25. Pales Espinosa E, Bouallegui Y, Grouzdev D, Brianik C, Czaja R, Geraci-Yee S, et al. An apicomplexan parasite drives the collapse of the bay scallop population in New York. Sci Rep. 2023;13:6655.
26. Mathur V, Kwong WK, Husnik F, Irwin NAT, Kristmundsson Á, Gestal C, et al. Phylogenomics identifies a new major subgroup of Apicomplexans, Marosporida *class nov.*, with extreme apicoplast genome reduction. Genome Biol Evol. 2021;13:evaa244.
27. Plough LV. Genetic load in marine animals: a review. Curr Zool. 2016;62:567–79.
28. Sutherland BJG, Itoh N, Gilchrist K, Boyle B, Roth M, Green TJ. Genomic diversity of wild and cultured Yesso scallop *Mizuhopecten yessoensis* from Japan and Canada. G3: Genes, Genomes, Genetics. 2023;13:jkad242.
29. Garcia-Gonzalo FR, Reiter JF. Scoring a backstage pass: mechanisms of ciliogenesis and ciliary access. J Cell Biol. 2012;197:697–709.
30. Bergmann C, Fliegauf M, Brüchle NO, Frank V, Olbrich H, Kirschner J, et al. Loss of nephrocystin-3 function can cause embryonic lethality, Meckel-Gruber-like syndrome, situs inversus, and renal-hepatic-pancreatic dysplasia. Am J Hum Genet. 2008;82:959–70.
31. Zhou W, Dai J, Attanasio M, Hildebrandt F. Nephrocystin-3 is required for ciliary function in zebrafish embryos. Am J Physiol Renal Physiol. 2010;299:F55–62.
32. Thi-Kim VuH, Rink JC, McKinney SA, McClain M, Lakshmanaperumal N, Alexander R, et al. Stem cells and fluid flow drive cyst formation in an invertebrate excretory organ. Elife. 2015;4: e07405.
33. Lin Z, Nie H, Zhang Y, Yin Z, Yan X. Genome-wide identification and analysis of HECT E3 ubiquitin ligase gene family in *Ruditapes philippinarum* and their involvement in the response to heat stress and Vibrio anguillarum infection. Comp Biochem Physiol D: Genomics Proteomics. 2022;43: 101012.
34. Cheng Q, Wang H, Jiang S, Wang L, Xin L, Liu C, et al. A novel ubiquitin-protein ligase E3 functions as a modulator of immune response against lipopolysaccharide in Pacific oyster. Crassostrea gigas Developmental & Comparative Immunology. 2016;60:180–90.
35. Mevissen TET, Komander D. Mechanisms of deubiquitinase specificity and regulation. Annu Rev Biochem. 2017;86:159–92.
36. Preta G, Cronin JG, Sheldon IM. Dynasore - not just a dynamin inhibitor. Cell Commun Signal. 2015;13:24.
37. Zhang J, Ferguson SS, Barak LS, Ménard L, Caron MG. Dynamin and beta-arrestin reveal distinct mechanisms for G protein-coupled receptor internalization. J Biol Chem. 1996;271:18302–5.
38. Mao J, Huang X, Sun H, Jin X, Guan W, Xie J, et al. Transcriptome analysis provides insight into adaptive mechanisms of scallops under environmental stress. Front Mar Sci. 2022;9: 971796.
39. Pales Espinosa E, Koller A, Allam B. Proteomic characterization of mucosal secretions in the eastern oyster, *Crassostrea virginica*. J Proteomics. 2016;132:63–76.
40. Liu G, Li Z, Yang M, Lin L, Liu J, Chen M. Functional characterization of a putative lipopolysaccharide-induced TNF-alpha factor (LITAF) from blood clam *Tegillarca granosa* in innate immunity. Fish Shellfish Immunol. 2020;97:390–402.
41. Lowry DB, Hoban S, Kelley JL, Lotterhos KE, Reed LK, Antolin MF, et al. Breaking RAD: an evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. Mol Ecol Resour. 2017;17:142–52.
42. Matz MV, Treml EA, Aglyamova GV, Bay LK. Potential and limits for rapid genetic adaptation to warming in a Great Barrier Reef coral. PLoS Genet. 2018;14: e1007220.
43. Reid BN, Star B, Pinsky ML. Detecting parallel polygenic adaptation to novel evolutionary pressure in wild populations: a case study in Atlantic cod (*Gadus morhua*). Phil Trans R Soc B. 2023;378:20220190.
44. Lachs L, Donner SD, Mumby PJ, Bythell JC, Humanes A, East HK, et al. Emergent increase in coral thermal tolerance reduces mass bleaching under climate change. Nat Commun. 2023;14:4939.
45. Farhat S, Tanguy A, Pales Espinosa E, Guo X, Boutet I, Smolowitz R, et al. Identification of variants associated with hard clam, *Mercenaria mercenaria*, resistance to Quahog Parasite Unknown disease. Genomics. 2020;112:4887–96.
46. Sambrook J, Fritsch EF, Maniatis T, Russell DW, Green MR. Molecular cloning: a laboratory manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1989.
47. Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. PLoS ONE. 2012;7: e37135.
48. Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. Stacks: an analysis tool set for population genomics. Mol Ecol. 2013;22:3124–40.
49. Rochette NC, Catchen JM. Deriving genotypes from RAD-seq short-read data using Stacks. Nat Protoc. 2017;12:2640–59.
50. Grouzdev D, Pales Espinosa E, Tettelbach S, Farhat S, Tanguy A, Boutet I, et al. Chromosome-level genome assembly of the bay scallop *Argopecten irradians*. Sci Data. 2024;11:1057.
51. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34:3094–100.
52. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25:2078–9.
53. O'Leary SJ, Puritz JB, Willis SC, Hollenbeck CM, Portnoy DS. These aren't the loci you're looking for: Principles of effective SNP filtering for molecular ecologists. Mol Ecol. 2018;27:3193–206.
54. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. Bioinformatics. 2011;27:2156–8.
55. Knaus BJ, Grünwald NJ. VCFR : a package to manipulate and visualize variant call format data in R. Mol Ecol Resour. 2017;17:44–53.
56. Hollenbeck CM, Portnoy DS, Garcia De La Serrana D, Magnesen T, Matejusova I, Johnston IA. Temperature-associated selection linked to putative chromosomal inversions in king scallop (*Pecten maximus*). Proc R Soc B. 2022;289:20221573.

57. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. Nat Genet. 2010;42:565–9.
58. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS Genet. 2006;2: e190.
59. Elhaik E. Principal Component Analyses (PCA)-based findings in population genetic studies are highly biased and must be reevaluated. Sci Rep. 2022;12:14683.
60. Yao Y, Ochoa A. Limitations of principal components in quantitative genetic association models for human studies. Elife. 2023;12: e79238.
61. Browning BL, Tian X, Zhou Y, Browning SR. Fast two-stage phasing of large-scale sequence data. The American Journal of Human Genetics. 2021;108:1880–90.
62. Privé F, Aschard H, Ziyatdinov A, Blum MGB. Efficient analysis of large-scale genome-wide data with two R packages: bigstatsr and bigsnpr. Bioinformatics. 2018;34:2781–7.
63. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–75.
64. Pembleton LW, Cogan NOI, Forster JW. StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. Mol Ecol Resour. 2013;13:946–52.
65. Raj A, Stephens M, Pritchard JK. fastSTRUCTURE: variational inference of population structure in large SNP data sets. Genetics. 2014;197:573–89.
66. Puechmaille SJ. The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. Mol Ecol Resour. 2016;16:608–27.
67. Li Y-L, Liu J-X. StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. Mol Ecol Resour. 2018;18:176–7.
68. Francis RM. POPHELPER : an R package and web app to analyse and visualize population structure. Mol Ecol Resour. 2017;17:27–32.
69. Malinsky M, Trucchi E, Lawson DJ, Falush D. RADpainter and fineRADstructure: Population Inference from RADseq Data. Mol Biol Evol. 2018;35:1284–90.
70. Whitlock MC, Lotterhos KE. Reliable Detection of Loci Responsible for Local Adaptation: Inference of a Null Model through Trimming the Distribution of F(ST). Am Nat. 2015;186(Suppl 1):S24–36.
71. Luu K, Bazin E, Blum MGB. pcadapt: an R package to perform genome scans for selection based on principal component analysis. Mol Ecol Resour. 2017;17:67–77.
72. De Jong MJ, Lovatt F, Hoelzel AR. Detecting genetic signals of selection in heavily bottlenecked reindeer populations by comparing parallel founder events. Mol Ecol. 2021;30:1642–58.
73. De Jong MJ, De Jong JF, Hoelzel AR, Janke A. SambaR: An R package for fast, easy and reproducible population-genetic analyses of biallelic SNP data sets. Mol Ecol Resour. 2021;21:1369–79.
74. Jombart T, Collins C. A tutorial for Discriminant Analysis of Principal Components (DAPC) using adegenet 2.0.0. London: Imperial College London, MRC Centre for Outbreak Analysis and Modelling; 2015.
75. Fisher RA. On the Interpretation of χ 2 from Contingency Tables, and the Calculation of P. J Roy Stat Soc. 1922;85:87.
76. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Stat Methodol. 1995;57:289–300.
77. Bonferroni C. Teoria statistica delle classi e calcolo delle probabilita. Pubblicazioni del R istituto superiore di scienze economiche e commericiali di firenze. 1936;8:3–62.

Publisher's Note