Parasitology

cambridge.org/par

Co-phylogeographic structure in a disease-causing parasite and its oyster host

Elizabeth Faye Weatherup[1,2], Ryan Carnegie[1], Allan E. Strand[3] and Erik E. Sotka[3]

[1]Virgina Institute of Marine Science, William & Mary, Gloucester Point, VA, USA; [2]Department of Biology and Marine Biology, University of North Carolina Wilmington, Wilmington, North Carolina, USA and [3]College of Charleston Marine Laboratory and Department of Biology, College of Charleston, Charleston, SC, USA

Research Article

Cite this article: Weatherup EF, Carnegie R, Strand AE, Sotka EE (2024). Co-phylogeographic structure in a disease-causing parasite and its oyster host. *Parasitology* **151**, 671–678. https://doi.org/10.1017/S0031182024000611

Received: 21 December 2023
Revised: 25 March 2024
Accepted: 12 May 2024
First published online: 21 May 2024

Keywords:
co-phylogeny; *Crassostrea virginica*; genetic divergence; genotype-by-sequencing; *Perkinsus marinus*

Corresponding author:
Elizabeth Faye Weatherup;
Email: efw4349@uncw.edu

Abstract

With the increasing affordability of next-generation sequencing technologies, genotype-by-sequencing has become a cost-effective tool for ecologists and conservation biologists to describe a species' evolutionary history. For host–parasite interactions, genotype-by-sequencing can allow the simultaneous examination of host and parasite genomes and can yield insight into co-evolutionary processes. The eastern oyster, *Crassostrea virginica*, is among the most important aquacultured species in the United States. Natural and farmed oyster populations can be heavily impacted by 'dermo' disease caused by an alveolate protist, *Perkinsus marinus*. Here, we used restricted site-associated DNA sequencing (RADseq) to simultaneously examine spatial population genetic structure of host and parasite. We analysed 393 single-nucleotide polymorphisms (SNPs) for *P. marinus* and 52,100 SNPs for *C. virginica* from 36 individual oysters from the Gulf of Mexico (GOM) and mid-Atlantic coastline. All analyses revealed statistically significant genetic differentiation between the GOM and mid-Atlantic coast populations for both *C. virginica* and *P. marinus*, and genetic divergence between Chesapeake Bay and the outer coast of Virginia for *C. virginica*, but not for *P. marinus*. A co-phylogenetic analysis confirmed significant coupled evolutionary change between host and parasite across large spatial scales. The strong genetic divergence between marine basins raises the possibility that oysters from either basin would not be well adapted to parasite genotypes and phenotypes from the other, which would argue for caution with regard to both oyster and parasite transfers between the Atlantic and GOM regions. More broadly, our results demonstrate the potential of RADseq to describe spatial patterns of genetic divergence consistent with coupled evolution.

Introduction

Microorganisms with obligate relationships with a host species commonly evolve in response to evolutionary changes in the host, by natural selection (e.g. host resistance or tolerance), or to parallel genetic divergence across biogeographic regions (Van Valen, 1973; Day, 1974; Ronquist, 1997; Page and Charleston, 1998). When host fitness is affected negatively by the presence of parasites, hosts may respond to evolutionary changes in the parasite, thus yielding co-evolutionary dynamics (Thompson, 2005). Co-evolutionary dynamics have significant implications for understanding and managing diseases in ecologically and economically important species (Coen and Bishop, 2015). Parasite tracking of host evolution, co-evolution or both yield significant patterns of coupled evolution between host and parasite genomes over space and time that can be quantified using emerging genotyping technologies (Vermeer *et al.*, 2011).

Most previous efforts to study coupled evolution of host–parasite interactions using genetics utilized microsatellites or Sanger sequencing of a 1 or a few loci, which is relatively time-consuming and expensive on a per-locus basis (Ebert and Fields, 2020; Märkle *et al.*, 2021). Recent advances in simultaneously amplifying single-nucleotide polymorphisms (or SNPs) using high-throughput sequencing technologies have allowed scientists the opportunity to use genotype-by-sequencing for both host and parasite genomes simultaneously. For example, Choi *et al.* (2014) used dual RNA sequencing to simultaneously look at *Brugia malayi* larvae and its mosquito host. Ansari *et al.* (2017) used a genome-wide association study (GWAS) of individuals infected with hepatitis C virus to show how SNPs in both the virus and host impacted the infection progression. Another study by Lees *et al.* (2019) also used GWAS to look at genomes of human hosts and *Streptococcus pneumoniae*, finding that host SNP variation is significant in differences in susceptibility to this bacterium. Most recently, Dexter *et al.* (2023) conducted a co-genome study of the planktonic crustacean, *Daphnia magna*, and its parasite *Pasteuria ramosa*, an endoparasitic bacterium (Dexter *et al.*, 2023). They found a signal of interspecies linkage disequilibrium across multiple sets of loci demonstrating the coevolution of this host and parasite system (Dexter *et al.*, 2023). Additionally, there have been several studies that have used restricted site-associated DNA sequencing (RADseq) to characterize host and parasite systems. Bracewell *et al.* (2018) used RADseq to find evidence of co-evolution and cascading speciation among 4 close interacting species, the western pine

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

beetle, *Dendroctonus brevicomis*, the beetle's mutualistic fungi, *Ceratocystiopsis brevicomi* and *Entomocorticium* sp., and the beetle's host tree, *Pinus ponderosa* (Bracewell *et al.*, 2018). Another study by Satler *et al.* (2019) was able to determine the frequent phenomena of host switching in Panamanian strangler figs and their pollinating fig wasps (Satler *et al.*, 2019). A similar occurrence was found in a study by Sweet *et al.* (2020). Sweet *et al.* used whole-genome sequencing and double-digested RADseq to obtain SNPs to examine co-evolutionary patterns and host-switching accessibility between 2 species of ptarmigan birds and their associated feather lice parasites, *Lagopoecus* and *Goniodes*. They found evidence of frequent host switching of lice among different bird populations in Alaska, as well as co-evolutionary patterns between the lice and birds (Sweet *et al.*, 2020).

The alveolate parasite *Perkinsus marinus* (Ph. Perkinsozoa) infects the eastern oyster *Crassostrea virginica* (Ph. Mollusca) along much of the oyster's geographic range from the Gulf of St. Lawrence, Canada to the Gulf of Mexico (GOM; Sparks, 1966). In addition to its ecological importance (Zimmerman *et al.*, 1989; Smaal and Prins, 1993), *C. virginica* has been the focus of historically as well as contemporarily significant fisheries and aquaculture industries. *Perkinsus marinus* causes 'dermo' disease, which can lead to widespread mortality in oyster populations (Carnegie *et al.*, 2021). Because of its significance, *P. marinus* has been a primary focus of regulation of aquaculture products among states of the US East Coast and the GOM, but the effectiveness of these regulations would benefit from deeper understanding of the ecological and evolutionary relationships between this host and parasite. To our knowledge, however, no previous genetic studies of *P. marinus* have genotyped SNPs nor simultaneously examined the oyster and parasite population genetics within the same individuals.

Unlike most other protozoans infecting molluscs, *P. marinus* can be cultured *in vitro*, which has made more analyses possible of the phenotypic and genetic structure of this parasite across its distribution than for other oyster parasites. In seminal early work, Bushek and Allen (1996) found that Atlantic isolates (VA and NJ) were more virulent than isolates from the GOM coast (TX and LA) in oysters collected from all 4 populations. Reece *et al.* (1997, 2001) genotyped restriction fragment length polymorphisms of *in vitro* cultures of *P. marinus*, noted the presence of diploid or multiple infections of *P. marinus* in single oysters and revealed 3 genetically distinct subdivisions among estuaries of the United States: the northeast Atlantic, southeast Atlantic and GOM (Reece *et al.*, 1997, 2001). Thompson *et al.* (2011, 2014) genotyped microsatellite markers and demonstrated the presence of dimorphic loci suggesting an ancient hybridization event that persists and a fairly complex, non-equilibria pattern of spatial population structure along the Atlantic and GOM coastlines (Thompson *et al.*, 2011, 2014).

Phylogeographic divergence among *C. virginica* populations is also well described. Multiple studies using a variety of mitochondrial and nuclear loci (Reeb and Avise, 1990; Hare and Avise, 1996; Varney *et al.*, 2009; Thongda *et al.*, 2018) have indicated divergence among and within the United States coastline of the GOM and Atlantic. These patterns reflect likely relatively short larval period (<2 weeks), entrainment of larvae within estuaries and possibly local purifying selection (Murray and Hare, 2006; Burford *et al.*, 2014).

Here, we used RADseq to examine SNPs of *P. marinus* and *C. virginica* within *P. marinus*-infected oysters. Our sampling was focused on infected adults collected at 2 GOM locations (Dauphin Island, AL, and Caminada Bay, LA) and Atlantic coast locations in Virginia (Great Wicomico River, Mockhorn Bay, York River, Burtons Bay, James River and Rappahannock River). We expected a parallel phylogeographic divergence in oyster and parasite genomes from the Gulf and Atlantic coasts yielding a significant co-phylogenetic signal partitioned across the geography of this host–parasite system.

Methods

Collection

We genotyped 95% ethanol-preserved oyster samples from Maryland (1 location), Virginia (14 locations), Alabama (1 location) and Louisiana (2 locations), as well as an *in vitro* culture of *P. marinus* originally collected in Galveston Bay, Texas, and 1 paraffin-embedded sample initially preserved in Davidson's fixative (Shaw and Battle, 1957) (Table 1). The Virginia samples were collected during routine surveillance for oyster diseases conducted by the Carnegie lab at the Virginia Institute of Marine Science (VIMS). Samples from other states represented aquaculture industry samples submitted for disease analysis in the context of interstate shellfish commerce. For each oyster sampled, animals were cleaned, measured and shucked to expose soft tissues. Gill and mantle were extracted from each oyster and preserved in 95% ethanol for genetic analysis.

Identification of oysters positive for Perkinsus marinus

DNA was extracted using a QIAGEN QIAamp DNA Mini Kit. Genomic DNA concentration was measured with a Nanodrop spectrophotometer to ensure the concentration was ≥100 ng μL^{-1}. If the concentration was >300 ng μL^{-1}, then the DNA was diluted in an equal volume of Milli-Q® water before polymerase chain reaction (PCR). A concentration of about 100 ng μL^{-1} was considered optimum for amplification of *P. marinus*-specific primers in each oyster sample. Samples were run on a PCR using PmarITS-70 forward primer (5′-CTT TTG YTW GAG WGT

Table 1. Oyster populations that were positive for *P. marinus* and extracted and sequenced [AL ($n = 3$), LA ($n = 9$), VA Bay ($n = 18$), VA Eastern ($n = 6$)].

Region, state and site	Total collected	Sequenced	Lat, long	Region	Oyster type
Atlantic VA Fleet Point, Great Wicomico	3	2	37.81, −76.29	VA Chesapeake Bay	Wild
Atlantic VA Oyster	6	5	37.28, −75.90	VA Eastern Shore seaside	Wild
Atlantic VA VIMS Beach	23	14	37.25, −76.50	VA Chesapeake Bay	Wild
Atlantic VA Wachapreague	1	1	37.61, −75.66	VA Eastern Shore seaside	Wild
Atlantic VA Wreck Shoal, James River	2	1	37.06, −76.29	VA Chesapeake Bay	Wild
Atlantic VA Broad Creek, Rappahannock River	1	1	37.58, −76.30	VA Chesapeake Bay	Wild
Gulf AL Auburn University	6	3	30.25, −88.08	Dauphin Island	Farmed
Gulf LA Caminada Bay	6	5	29.24, −90.00	Grand Isle	Farmed

https://doi.org/10.1017/S0031182024000611 Published online by Cambridge University Press

TGC GAG ATG-3′) and PmarITS-600 reverse primer (5′-CGA GTT TGC GAG TAC CTC KAG AG-3′), which target *P. marinus*. Denaturing was done at 94 °C for 30 s, annealing at 57 °C for 30 s and extension at 72 °C for 1.5 min. These steps were repeated 40 times. After PCR, to identify oysters that were positive for *P. marinus*, the PCR products were run on a 2% agarose gel at 100 volts for 30 min and stained with GelRed Nucleic Acid stain in a water bath for 25 min. This included a 1KB bp ladder at the first lane of the gel so that the desired 1000 base pair size could be visualized. Positive samples were identified and used for further analysis. The goal was to obtain 30 positive samples for each site, which was possible in some cases, but not all.

All positive samples in microcentrifuge tubes were randomly mixed before transport in case there was an error in future sequencing steps that could cause a loss in representation of an entire population. Once the microcentrifuge tubes containing the extracted DNA were thoroughly mixed, 20 μL of each of the 64 samples were used in subsequent library construction.

Library preparation

We prepared genomic libraries using the genotype-by-sequencing approach outlined in Parchman *et al.* (2012), applied to samples identified as *P. marinus*-positive. The DNA of each individual oyster was digested separately with 2 restriction enzymes, EcoRI and MseI. The digested DNA fragments were then ligated to Illumina adaptors at the MseI end and with Illumina adaptors coupled with an 8–10 bp unique barcode to the EcoRI end to allow identification of the individual *in silico*. The restriction-ligation products were then PCR-amplified in 2 separate reactions using standard Illumina primers. The final PCR products were pooled and shipped for sequencing.

Fifty-nine *P. marinus*-positive samples were sequenced in 2020 at the Tufts University Genomic Services (200–400 bp fraction; single-end sequencing on an Illumina HiSeq 2500 using SE125), 42 samples were sequenced in 2022 at the University of Texas Genomic Sequencing and Analysis Facility (300–450 bp fraction; single-end sequencing on an Illumina NovaSeq SP using SR100) and 37 of each of those samples were sequenced in both runs.

Data filtering

Reads were aligned to the *P. marinus* genome (GCA_000006405.1) using *bwa mem* (Li and Durbin, 2010) and *samtools/bcftools* 1.9 (Danecek *et al.*, 2021) using default settings. We also aligned these samples to the *C. virginica* genome (GCF_002022765.2). The median number of reads aligned to the *P. marinus* genome (0.13 M; a range of 0.10–1.01 M) was 0.37% of the median number aligned to the *C. virginica* genome (34 M; a range of 3.52–66.89 M). This yielded a ratio of 265 *C. virginica* to *P. marinus* reads per sample. Of the total number of reads per sample, a median of 0.80 and 91.30% of reads were *P. marinus* and *C. virginica*, respectively.

After alignment of 64 samples to the *P. marinus* genome, we kept 35 samples that had between 100 K and 1 M *P. marinus* reads. Another sample identified as 6837–16, from VA had 1.1 M reads, from which we randomly downsampled 500 K reads using samtools view. A final sample of 1993 TX live culture (TX-2-69, purchased from the American Type Culture Collection, ATCC) had 25 million reads that aligned with *P. marinus*; however, genotype calls showed a predominance of no calls (NAs) and consequently, it was removed. One likely reason for NAs at this step is that this culture contains multiple strains that interfere with genotype calling algorithms. For the remainder of the samples, we did not detect the presence of multiple strains nor patterns of dimorphic loci (Thompson *et al.*, 2014). Thus, we were left with 36 samples (Table 1).

For *P. marinus*, we used bcftools to find all SNPs [set minor allele frequency (MAF) threshold at 0] and then a custom script to find SNPs that had at least 1 read across 50% of individuals. This yielded 772 SNPs. An analysis of allele frequencies using angsd (-doMAF 1) indicated that 393 SNPs were polymorphic between or within samples. Thus, all subsequent analyses are based on samtools-generated phred-scale genotype likelihoods of 393 polymorphic SNPs at 36 individuals. For *C. virginica*, we created genotypes from the same 36 individuals, and included SNPs with MAF > 1% and at least 1 read per sample, yielding samtools-generated genotypes at 52 100 SNPs.

Principal component analysis

Principal component analysis (PCA) plots were generated for both organisms between collections. We used PCangsd (Meisner and Albrechtsen, 2018) on *P. marinus* genotype likelihoods and *prcomp* on *C. virginica* genotype calls within R (R Core Team, 2021). PCA is an exploratory analysis for large datasets providing more comprehensible information by grouping the data based on the individual samples' genotypic similarities.

Analysis of molecular variance

To assess how much genetic variation is partitioned among and within groups of samples, we performed hierarchical analysis of molecular variance (AMOVA; Excoffier *et al.*, 1992) using Pegas (Paradis, 2010). *P* values were generated using 1000 bootstrap replicates. For *P. marinus*, we converted the phred-scale genotype likelihoods per SNP-sample combination into probabilities that summed to 1 and then converted these to a single value ranging from 0 to 2, where 0, 1 and 2 represent the highest probability of a homozygote, heterozygote and alternative homozygote, respectively. This matrix served as an input to Pegas. For *C. virginica*, the input was genotype calls from samtools pulled using R:vcfR (Knaus and Grünwald, 2017) from a vcf file. Given the patterns found in PCA, we analysed 4 groups of samples: Alabama ($n = 3$), Louisiana ($n = 9$), Virginia Chesapeake Bay ($n = 18$) and Virginia Eastern Shore seaside ($n = 6$; see Table 1). We estimated expected heterozygosity in *P. marinus* using genotype likelihoods as implemented in *angsd*, and of genotype calls in *C. virginica* using *strataG* (Archer *et al.*, 2016).

Admixture analysis

We implemented maximum-likelihood admixture methods to infer ancestry of individuals. We used NGSadmix (Meisner and Albrechtsen, 2018) to analyse genotype likelihoods of *P. marinus* and LEA: SNMF (Frichot and François, 2015) to analyse genotype calls of *C. virginica*. Analyses were run across an *a priori* range of genetic clusters ($k = 2$–10) and replicated 10 times. Given the low number of samples, we only visualized up to $k = 6$ for both datasets. An analysis of logL changes (following Evanno *et al.*, 2005) indicates that the best-fit k to the *P. marinus* dataset was $k = 5$. The strongest cross-validation in the SNMF *C. virginica* analysis was $k = 6$.

Co-phylogenetic analysis

We visualized the co-phylogenetic structure between host and parasite using maximum-likelihood trees. We first generated fasta-formatted files of concatenated SNPs coded as IUPAC nucleotides using seqinr (Charif and Lobry, 2007). For *P. marinus*, we used PhyML implemented in SeaView (Gouy *et al.*, 2010) with default parameters (GTR + 4 rate classes as a model of evolution; 100 bootstrap replicates). For *C. virginica*, we used IQTree (Trifinopoulos *et al.*, 2016), which had

https://doi.org/10.1017/S0031182024000611 Published online by Cambridge University Press

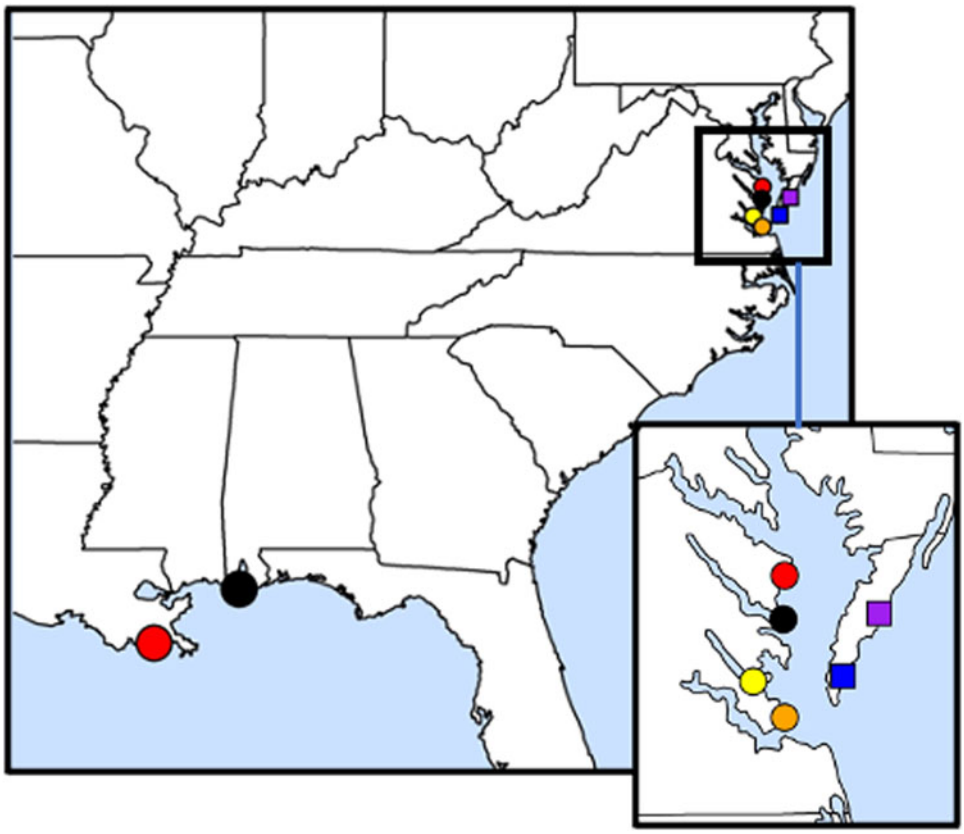


Figure 1. Map of geographic location of each site collection for the region of Virginia [circle, Chesapeake Bay; squares, Eastern Shore; red, Fleet Point ($n = 2$); yellow, VIMS Beach ($n = 11$); orange, Wreck Shoal ($n = 1$); black, Broad Creek ($n = 1$); purple, Wachapreague ($n = 1$); blue, Oyster ($n = 6$)] and 2 regions in the Gulf of Mexico [red = Louisiana ($n = 5$) and black = Alabama ($n = 3$)].

12766 parsimony-informative and 10028 singleton SNPs. The best-fit model of evolution was TVM + F + I + G4, allele frequencies were computed from the alignment and we used the ultrafast method to create 1000 bootstrap replicates. To statistically assess cophylogenetic signal among host and parasite, we used these trees as input to the Procrustean Approach to Co-phylogeny, or R::paco (Hutchinson *et al.*, 2017). This approach assumes dependence of the parasite phylogeny on host, and assesses the probability that the observed network has more phylogenetic congruence than 1000 random instances of the interaction network (Fig. 1).

Results

Strong genetic divergence between the GOM collections (LA and AL) and Virginia were revealed within the nuclear genomes of both *P. marinus* and *C. virginica* in all analyses. PCA (Fig. 2), admixture analyses (Fig. 3) and an AMOVA (Table 2) revealed strong divergence in *P. marinus* and *C. virginica* genotypes.

In contrast, there was discordance between host and parasite in geographic divergence at smaller spatial scales within Virginia. *Crassostrea virginica* genotypes between Chesapeake Bay and seaside Eastern Shore groups were divergent in PCA (Fig. 2B), admixture (Fig. 3B) and AMOVA (Table 2). However, *P. marinus* genotypes were not clearly divergent in any of these analyses (Figs 2A, 3A and 4; Table 2). Regions also did not differ in expected heterozygosity for either *P. marinus* or *C. virginica* (Table 3).

A co-phylogeny plot (Fig. 4) provides a view of both the phylogeographic structure and patterns of mixing between host and parasite genotypes. Overall, paco (Hutchinson *et al.*, 2017)

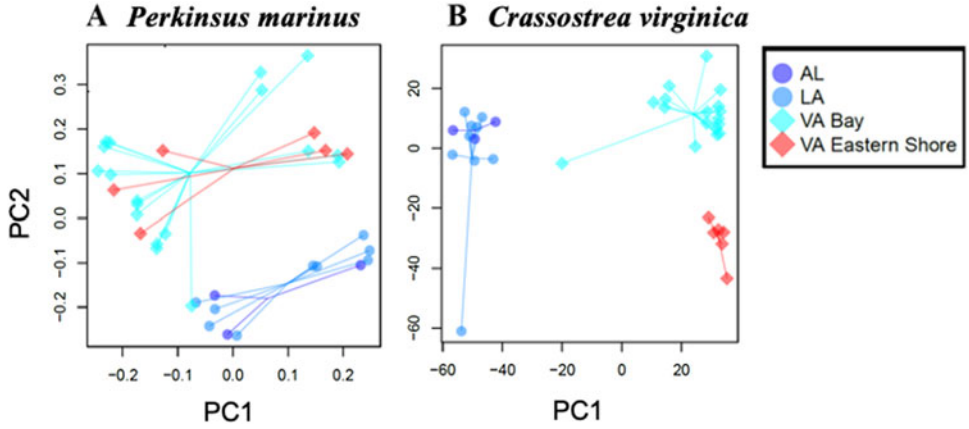


Figure 2. Principal components analyses of (A) *Perkinsus marinus* (393 loci, $n = 36$) from PCangsd and (B) *Crassostrea virginica* (52100 SNPs, $n = 35$) from prcomp.

https://doi.org/10.1017/S0031182024000611 Published online by Cambridge University Press

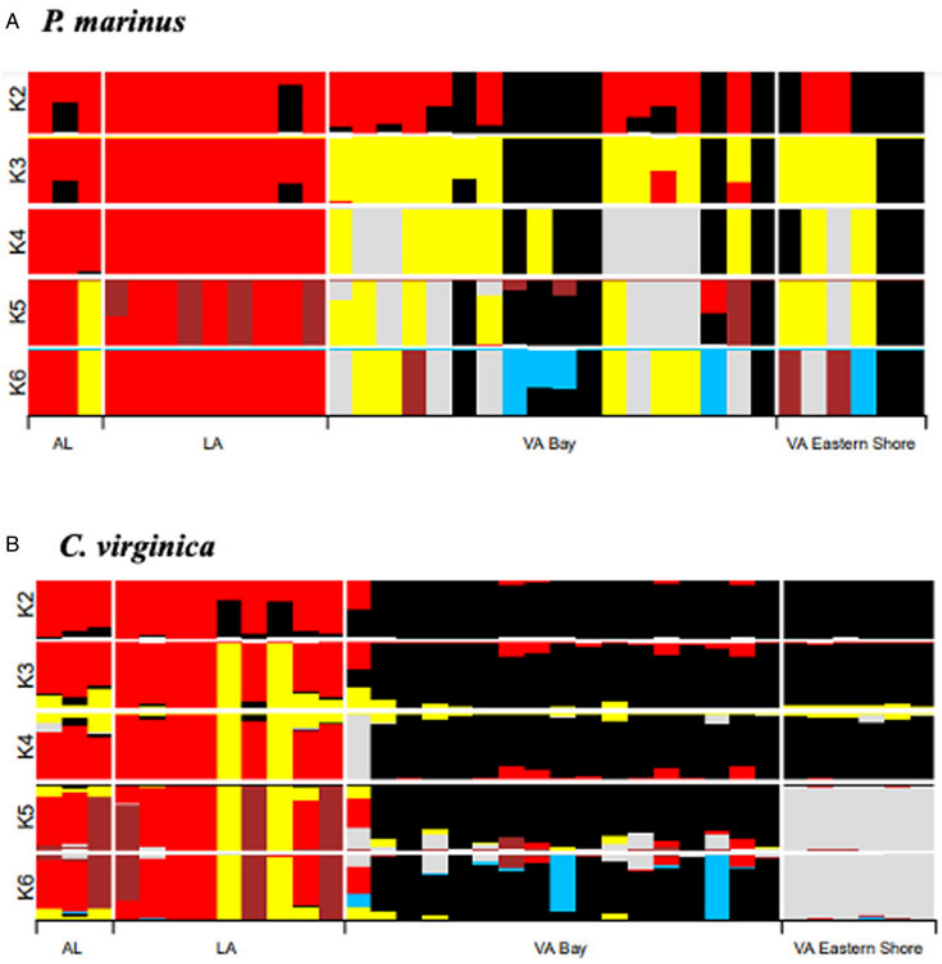


Figure 3. Admixture analysis of (A) *Perkinsus marinus* (393 loci, $n = 36$) using genotype likelihoods in NGSadmix and (B) *Crassostrea virginica* (52100 SNPs, $n = 35$) using genotypes in SNMF.

indicated a significant co-phylogenetic signal between the groups (m2xy = 0.180, $P < 0.001$, $n = 1000$). Thus, across all samples, the parasite phylogeny significantly tracked the oyster phylogeny, indicating they have undergone coupled evolutionary change, at least at the larger GOM *vs* Atlantic spatial scale. There were 2 oyster samples collected in GOM (coded LA6946-C and AL7040-6) that had *P. marinus* genotypes that were embedded within the Virginia clade (Fig. 4). This may have reflected ancestral polymorphisms from a single ancestor that have not been sorted, poor phylogenetic support in the ML tree of the *P. marinus* or more recent movement of *P. marinus* from the Atlantic to the GOM.

Table 2. Analysis of molecular variance (AMOVA) on *Perkinsus marinus* and *Crassostrea virginica* samples

P. marinus	AL	LA	VA Bay	VA Eastern
AL		0.041	0.035	0.013
LA	0.146		0.006	0.002
VA Bay	0.208	0.113		0.133
VA Eastern	0.294	0.53	0.047	
C. virginica	AL	LA	VA Bay	VA Eastern
AL		0.681	0.001	0.007
LA	−0.028		<0.001	<0.001
VA Bay	0.165	0.211		<0.001
VA Eastern	0.238	0.281	0.038	

Overall PhiST (i.e. across 4 populations) was 0.125 ($P = 0.005$) and 0.165 ($P < 0.001$) for *P. marinus* and *C. virginica*, respectively. Pairwise-PhiST values are shown below and above diagonals, respectively.

Discussion

In our analyses, clear genotypic differentiation was evident between our Atlantic samples and those from the GOM, in both *P. marinus* and *C. virginica* (Figs 2 and 3; Table 2). Furthermore, we identified genotypic differentiation within *C. virginica* populations between the Chesapeake Bay and the seaside Eastern Shore in Virginia (Figs 2 and 3; Table 2). These findings support previous research on *P. marinus* (Reece *et al.*, 1997, 2001; Thompson *et al.*, 2014) that documented the genetic differentiation of *P. marinus* between the GOM and Mid-Atlantic regions. It is worth noting that expanding our sampling of individuals and loci could potentially unveil significant genetic structure of *P. marinus* populations within regions. Our results also align with the conclusions of Varney *et al.* (2009) and Thongda *et al.* (2018) for *C. virginica*, of genetic differentiation between and within regions. This strong genetic structure within regions likely reflects barriers to larval dispersal, local selection or both (Stauber, 1950; Narváez *et al.*, 2012).

Our maximum-likelihood analysis (Fig. 4) revealed a strong signal of co-phylogeny between the parasite *P. marinus* and its oyster host. A close interaction between these 2, documented for nearly three-quarters of a century but potentially of far older origins, has resulted in an ongoing arms race, as these

https://doi.org/10.1017/S0031182024000611 Published online by Cambridge University Press

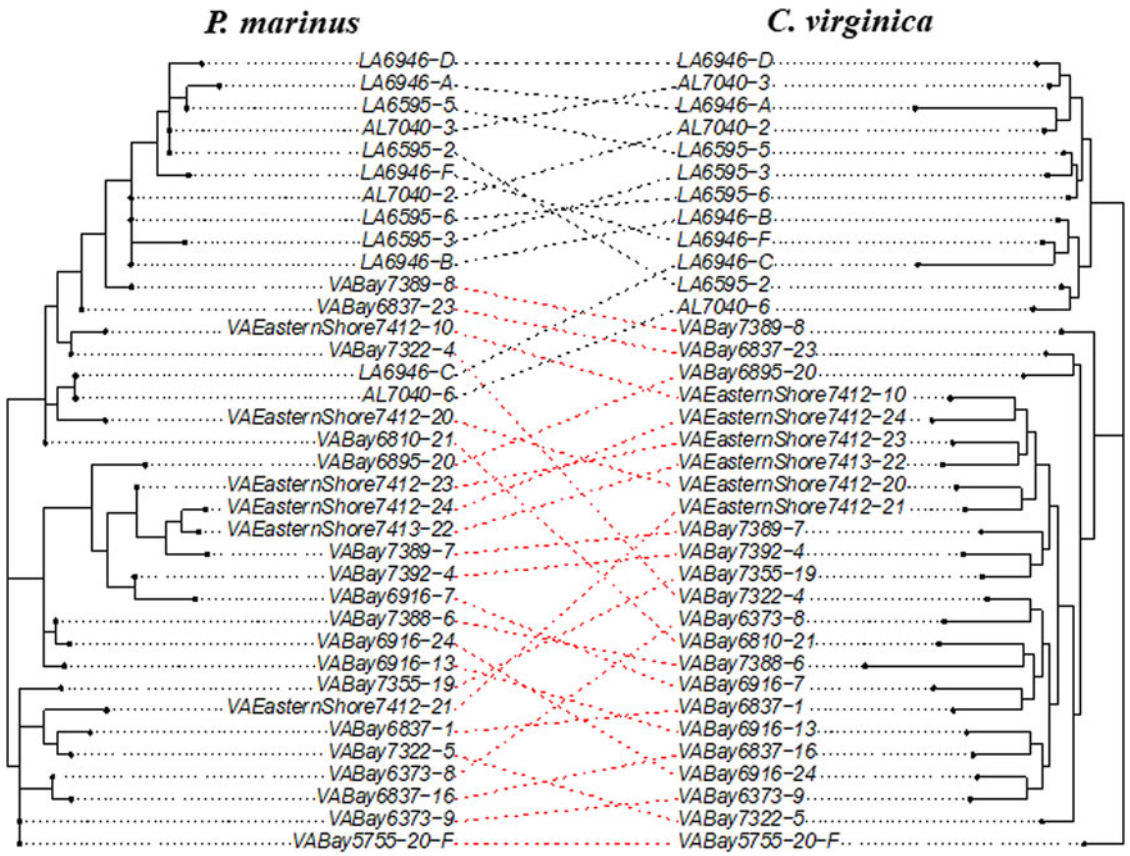


Figure 4. A maximum-likelihood co-phylogeny of the parasite *Perkinsus marinus* (392 bp) and host *Crassostrea virginica* (52052 bp). All nodes have 100% consensus support for *C. virginica* while all nodes have <50% support for *P. marinus* (1000 bootstrap replicates for both). Black and red dashed lines indicate GOM and VA genotypes within *Crassostrea*, respectively, and linked to *Perkinsus* genotypes.

organisms are continually adapting to outcompete one another. This phenomenon is observed in numerous other species (Dismukes *et al.*, 2022), and this methodology has the potential to uncover a multitude of evolutionary histories between hosts and parasites across diverse species.

This approach enabled us to detect genetic structure between the GOM and Atlantic *P. marinus* populations, which underscores its potential effectiveness to resolve the regional genetic diversity more finely in this pathogen. The implications for management of *P. marinus* in the context of regional shellfish aquaculture commerce are already important, even if this initial analysis was too limited to provide clear perspective on intra-regional diversity. Resource managers in the states affected by *P. marinus* parasitism of oysters are keenly interested in the question of genetic structure, because of the possibility of inadvertently creating new encounters, by pathogen introduction from distant waters through aquaculture transfers, between local oysters and *P. marinus* 'strains', or phenotypes, to which local oyster populations are not well adapted. The concern is that this would potentially result in a *P. marinus* epizootic of increased severity and economic destruction. Even if *P. marinus* is widely distributed and widely similar in prevalence across locations, which would argue that any aquaculture transfer of oysters with a low prevalence of *P. marinus* infections should be inconsequential against the natural backdrop of highly prevalent *P. marinus*, there is an increasing reluctance to risk introductions of *P. marinus* along with shellfish transfers because of genetic concerns. Better resolution of *P. marinus* genetic (and ideally, phenotypic) diversity could allow determination of 'zones' of common genotypic and phenotypic profiles within which control of *P. marinus* in transfers might be relaxed, to the benefit of reasonable aquaculture commerce and aquaculture biosecurity generally (Bushek and Allen, 1996; Reece *et al.*, 1997, 2001; Carnegie *et al.*, 2016). Based on our analyses, a case could already be made that divergence between Atlantic and GOM *P. marinus* populations would suggest that they should be precautionarily managed as distinct 'zones', between which transfers of allopatric parasite as well as oyster genotypes and phenotypes should be avoided.

This study also resulted in analysis among 393 SNPs for the parasite, despite the samples being dominated by host DNA, as

Table 3. Expected heterozygosity (mean ± standard error across individuals) for *Crassostrea virginica* and *Perkinsus marinus*

	n	He *C. virginica*	s.e. *C. virginica*	He *P. marinus*	s.e. *P. marinus*
AL	3	0.294	0.155	0.015	0.046
LA	9	0.297	0.079	0.016	0.026
VA Bay	18	0.286	0.046	0.029	0.024
VA Eastern Shore	6	0.281	0.089	0.032	0.046

https://doi.org/10.1017/S0031182024000611 Published online by Cambridge University Press

it was taken from a gill and mantle sample from each oyster. The analyses were able to be examined at 52 100 SNPs for the host genome, *C. virginica*, demonstrating that with a high alignment of reads to the genome, this method can yield deep sequencing analysis, which is necessary to providing fine resolution of population genetics of species to aid in disease and risk management (Bernatchez *et al*., 2019).

Future steps

The process of uncovering multiple loci simultaneously for both host and parasite species through RADseq and Illumina sequencing holds immense promise for future population genetic studies. However, it is important to acknowledge the limitations of our analysis, primarily stemming from the relatively small sample size and limited sample locations. To enhance the utility of this tool for co-phylogeny studies, several steps can be taken, especially regarding the parasite analysis.

First, expanding the sample size from each location to include a minimum of 30 or more *P. marinus*-positive oysters would strengthen the robustness of our analyses. This larger dataset would provide a more comprehensive understanding of the genetic structure of the parasite within and across regions, potentially allowing us to identify unique genotypes with important implications for management strategies. Another improvement involves finding a method to enrich parasite DNA, as it is significantly overshadowed by host DNA. This could be achieved by incorporating a pre-amplification step utilizing beads with attached oligonucleotides designed to specifically target the desired parasite loci, a technique supported by prior research (Shapero *et al*., 2001; Rödiger *et al*., 2014). Through the application of such amplification methods and an increased sample size, we could gain deeper insights into the genetic structure of the parasite.

Conclusion

This study highlights genotypic divergence among regions and within regions of *P. marinus* and its host, *C. virginica* using RADseq. Unlike other studies (Bracewell *et al*., 2018; Satler *et al*., 2019; Sweet *et al*., 2020) that used RADseq, this study amplified genomes of both host and parasites from a naturally infected host, rather than a lab infection or separate sequencing. This allowed us to get an accurate and current view on the evolutionary interactions in this host and parasite system. Moreover, RADseq offers a straightforward means of assessing coupled evolutionary change within host and parasite species simultaneously.

This study could be improved with an increased sampling size of wild oysters, more sampling sites and a method to increase parasite loci. However, despite the low sampling size, this methodology provided us insight into this host parasite interaction effectively.

Data availability statement. Relevant code and datasets are found at https://github.com/esotka/WeatherupPerkinsus. FASTQ files have been uploaded to GenBank (BioSample accession SAMN39856659–95).

Author contributions. All authors conceived and designed the study and edited the final manuscript. R. C. and E. F. W. collected samples, and E. F. W. extracted samples. E. F. W. and E. E. S. prepared the samples for sequencing, performed all analyses, generated figures and tables and wrote the manuscript.

Financial support. This research was funded by the National Science Foundation (OCE-1924599) and the VIMS Foundation A. Marshall Acuff, Sr., Memorial Endowment for Oyster Disease Research.

Competing interests. None.

Ethical standards. Not applicable.

References

Ansari MA, Pedergnana V, Ip LC, Magri A, Von Delft A, Bonsall D, Chaturvedi N, Bartha I, Smith D, Nicholson G, McVean G, Trebes A, Piazza P, Fellay J, Cooke G, Foster GR, Hudson E, McLauchlan J, Simmonds P, Bowden R, Klenerman P, Barnes E and Spencer CCA (2017) Genome-to-genome analysis highlights the effect of the human innate and adaptive immune systems on the hepatitis C virus. *Nature Genetics* **49**, 666–673.

Archer FI, Adams PE, and Schneiders BB (2017) stratag: An r package for manipulating, summarizing and analysing population genetic data. *Molecular Ecology Resources* **17**, 5–11.

Bernatchez S, Xuereb A, Laporte M, Benestan L, Steeves R, Laflamme M, Bernatchez L and Mallet M (2019) Seascape genomics of eastern oyster (*Crassostrea virginica*) along the Atlantic coast of Canada. *Evolutionary Applications* **12**, 587–609.

Bracewell RR, Vanderpool D, Good JM and Six DL (2018) Cascading speciation among mutualists and antagonists in a tree–beetle–fungi interaction. *Proceedings of the Royal Society B: Biological Sciences* **285**, 20180694.

Burford MO, Scarpa J, Cook BJ and Hare MP (2014) Local adaptation of a marine invertebrate with a high dispersal potential: evidence from a reciprocal transplant experiment of the eastern oyster *Crassostrea virginica*. *Marine Ecology Progress Series* **505**, 161–175.

Bushek D and Allen S (1996) Host-parasite interactions among broadly distributed populations of the eastern oyster *Crassostrea virginica* and the protozoan *Perkinsus marinus*. *Marine Ecology Progress Series* **139**, 127–141.

Carnegie RB, Arzul I and Bushek D (2016) Managing marine diseases in the context of regional and international commerce: policy issues and emerging concerns. *Philosophical Transactions of the Royal Society B* **371**, 20150215.

Carnegie RB, Ford SE, Crockett RK, Kingsley-Smith PR, Bienlien LM, Safi LSL, Whitefleet-Smith LA and Burreson EM (2021) A rapid phenotype change in the pathogen *Perkinsus marinus* was associated with a historically significant marine disease emergence in the eastern oyster. *Scientific Reports* **11**, 12872.

Charif D and Lobry JR (2007) SeqinR 1.0-2: a contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. In Bastolla U, Porto M, Roman HE and Vendruscolo M (eds), *Structural Approaches to Sequence Evolution: Molecules, Networks, Populations*. Berlin, Heidelberg: Springer, pp. 207–232. doi: 10.1007/978-3-540-35306-5_10

Choi Y-J, Aliota MT, Mayhew GF, Erickson SM and Christensen BM (2014) Dual RNA-seq of parasite and host reveals gene expression dynamics during filarial worm–mosquito interactions. *PLoS Neglected Tropical Diseases* **8**, e2905.

Coen LD and Bishop MJ (2015) The ecology, evolution, impacts and management of host-parasite interactions of marine molluscs. *Journal of Invertebrate Pathology* **131**, 177–211.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM and Li H (2021) Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008.

Day PR (1974) *Genetics of Host-Parasite Interaction*. San Francisco: WH Freeman and Co.

Dexter E, Fields PD and Ebert D (2023) Uncovering the genomic basis of infection through co-genomic sequencing of hosts and parasites. *Molecular Biology and Evolution* **40**, msad145.

Dismukes W, Braga MP, Hembry DH, Heath TA and Landis MJ (2022) Cophylogenetic methods to untangle the evolutionary history of ecological interactions. *Annual Review of Ecology, Evolution, and Systematics* **53**, 275–298.

Ebert D and Fields PD (2020) Host-parasite co-evolution and its genomic signature. *Nature Reviews. Genetics* **21**, 754–768.

Evanno G, Regnaut S and Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611–2620.

Excoffier L, Smouse PE and Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, 479–491.

Frichot E and François O (2015) LEA: an R package for landscape and ecological association studies. *Methods in Ecology and Evolution* **6**, 925–929.

Gouy M, Guindon S and Gascuel O (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution* **27**, 221–224.

https://doi.org/10.1017/S0031182024000611 Published online by Cambridge University Press

Hare MP and Avise JC (1996) Molecular genetic analysis of a stepped multilocus cline in the American oyster (*Crassostrea virginica*). *Evolution; International Journal of Organic Evolution* **50**, 2305–2315.

Hutchinson MC, Cagua EF, Balbuena JA, Stouffer DB and Poisot T (2017) paco: Implementing procrustean approach to cophylogeny in R. *Methods in Ecology and Evolution* **8**, 932–940.

Knaus BJ and Grünwald NJ (2017) vcfr: A package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources* **17**, 44–53.

Lees JA, Ferwerda B, Kremer PHC, Wheeler NE, Serón MV, Croucher NJ, Gladstone RA, Bootsma HJ, Rots NY, Wijmega-Monsuur AJ, Sanders EAM, Trzciński K, Wyllie AL, Zwinderman AH, van den Berg LH, van Rheenen W, Veldink JH, Harboe ZB, Lundbo LF, de Groot LCPGM, van Schoor NM, van der Velde N, Ängquist LH, Sørensen TIA, Nohr EA, Mentzer AJ, Mills TC, Knight JC, du Plessis M, Nzenze S, Weiser JN, Parkhill J, Madhi S, Benfield T, von Gottberg A, van der Ende A, Brouwer MC, Barrett JC, Bentley SD and van de Beek D (2019) Joint sequencing of human and pathogen genomes reveals the genetics of pneumococcal meningitis. *Nature Communications* **10**, 2176.

Li H and Durbin R (2010) Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595.

Märkle H, John S, Cornille A, Fields PD and Tellier A (2021) Novel genomic approaches to study antagonistic coevolution between hosts and parasites. *Molecular Ecology* **30**, 3660–3676.

Meisner J and Albrechtsen A (2018) Inferring population structure and admixture proportions in low-depth NGS data. *Genetics* **210**, 719–731.

Murray M and Hare M (2006) A genomic scan for divergent selection in a secondary contact zone between Atlantic and Gulf of Mexico oysters, *Crassostrea virginica*. *Molecular Ecology* **15**, 4229–4242.

Narváez DA, Klinck JM, Powell EN, Hofmann EE, Wilkin J and Haidvogel DB (2012) Modeling the dispersal of eastern oyster (*Crassostrea virginica*) larvae in Delaware Bay. *Journal of Marine Research* **70**, 381–409.

Page RD and Charleston MA (1998) Trees within trees: phylogeny and historical associations. *Trends in Ecology & Evolution* **13**, 356–359.

Paradis E (2010) pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics* **26**, 419–420.

Parchman TL, Gompert Z, Mudge J, Schilkey FD, Benkman CW and Buerkle CA (2012) Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology* **21**, 2991–3005.

R Core Team (2021) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at https://www.R-project.org/

Reeb CA and Avise JC (1990) A genetic discontinuity in a continuously distributed species: mitochondrial DNA in the American oyster, *Crassostrea virginica*. *Genetics* **124**, 397–406.

Reece KS, Bushek D and Graves JE (1997) Molecular markers for population genetic analysis of *Perkinsus marinus*. *Molecular Marine Biology and Biotechnology* **6**, 197–206.

Reece KS, Bushek D, Hudson KE and Graves JE (2001) Geographic distribution of *Perkinsus marinus* genetic strains along the Atlantic and Gulf Coasts of the USA. *Marine Biology* **139**, 1047–1055.

Rödiger S, Liebsch C, Schmidt C, Lehmann W, Resch-Genger U, Schedler U and Schierack P (2014) Nucleic acid detection based on the use of microbeads: a review. *Microchimica Acta* **181**, 1151–1168.

Ronquist F (1997) Phylogenetic approaches in coevolution and biogeography. *Zoologica Scripta* **26**, 313–322.

Satler JD, Herre EA, Jandér KC, Eaton DAR, Machado CA, Heath TA and Nason JD (2019) Inferring processes of coevolutionary diversification in a community of Panamanian strangler figs and associated pollinating wasps*. *Evolution* **73**, 2295–2311.

Shapero MH, Leuther KK, Nguyen A, Scott M and Jones KW (2001) SNP genotyping by multiplexed solid-phase amplification and fluorescent minisequencing. *Genome Research* **11**, 1926–1934.

Shaw BL and Battle HI (1957) The gross and microscopic anatomy of the digestive tract of the oyster Crassostrea virginica (Gmelin). *Canadian Journal of Zoology* **35**, 325–347.

Smaal AC and Prins TC (1993) The uptake of organic matter and the release of inorganic nutrients by bivalve suspension feeder beds. In Dame RF (ed.), *Bivalve Filter Feeders*. Berlin, Heidelberg: Springer, pp. 271–298. doi: 10.1007/978-3-642-78353-1_8.

Sparks AK (1966) GALTSOFF, P. S. 1964. The American oyster *Crassostrea virginica* Gmelin. Fishery Bulletin, v. 64. United States Government Printing Office, Washington, D. C. iii + 480 p, $2.75. *Limnology and Oceanography* **11**, 312–312.

Stauber LA (1950) The problem of physiological species with special reference to oysters and oyster drills. *Ecology* **31**, 109–118.

Sweet AD, Wilson RE, Sonsthagen SA and Johnson KP (2020) Lousy grouse: comparing evolutionary patterns in Alaska galliform lice to understand host evolution and host–parasite interactions. *Ecology and Evolution* **10**, 8379–8393.

Thompson JN (2005) *The Geographic Mosaic of Coevolution*. Chicago, IL: University of Chicago Press.

Thompson PC, Rosenthal BM and Hare MP (2011) An evolutionary legacy of sex and clonal reproduction in the protistan oyster parasite *Perkinsus marinus*. *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* **11**, 598–609.

Thompson P, Rosenthal B and Hare M (2014) Microsatellite genotypes reveal some long-distance gene flow in *Perkinsus marinus*, a major pathogen of the eastern oyster, *Crassostrea virginica* (Gmelin). *Journal of Shellfish Research* **33**, 195–206.

Thongda W, Zhao H, Zhang D, Jescovitch LN, Liu M, Guo X, Schrandt M, Powers SP and Peatman E (2018) Development of SNP panels as a new tool to assess the genetic diversity, population structure, and parentage analysis of the eastern oyster (*Crassostrea virginica*). *Marine Biotechnology* **20**, 385–395.

Trifinopoulos J, Nguyen L-T, von Haeseler A and Minh BQ (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research* **44**, W232–W235.

Van Valen L (1973) A new evolutionary law. *Evolution* **1**, 1–30.

Varney RL, Galindo-Sánchez CE, Cruz P and Gaffney PM (2009) Population genetics of the eastern oyster *Crassostrea virginica* (Gmelin, 1791) in the Gulf of Mexico. *Journal of Shellfish Research* **28**, 855–864.

Vermeer KMCA, Dicke M and De Jong PW (2011) The potential of a population genomics approach to analyze geographic mosaics of plant – insect coevolution. *Evolutionary Ecology* **25**, 977–992.

Zimmerman RJ, Minello TJ, Baumer T and Castiglione MC (1989) Oyster reef as habitat for estuarine macrofauna. *National Oceanic and Atmospheric Administration Technical Memorandum* NMFSSEFC-249.

https://doi.org/10.1017/S0031182024000611 Published online by Cambridge University Press