This is an Accepted Manuscript for Parasitology. This version may be subject to change during the production process. DOI: 10.1017/S0031182024000611

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

E. F. Weatherup[1,2], R.B. Carnegie[1], A. E. Strand[3], E. E. Sotka[3]

[1]Virgina Institute of Marine Science, William & Mary, P.O. Box 1346, Gloucester Point, VA, USA

[2]Present address: Department of Biology and Marine Biology, University of North Carolina Wilmington, Wilmington, North Carolina, USA

[3]College of Charleston Marine Laboratory and Department of Biology, College of Charleston, 205 Fort Johnson Road, Charleston SC 29412, USA

Corresponding author: Elizabeth Weatherup, Email: efw4349@uncw.edu and Erik Sotka, Email: SotkaE@cofc.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 analyzed 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 Gulf of Mexico regions. More broadly, our results demonstrate the potential of RADseq to describe spatial patterns of genetic divergence consistent with coupled evolution.

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

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

Key Findings

- Simultaneous genotyping of SNPs within a host oyster and parasitic alveolate using RADseq indicate strong genetic differentiation at large spatial scales indicating significant coupled evolutionary change.
- Statistically significant genetic divergence between Chesapeake Bay and the outer coast of Virginia was detected for *C. virginica* but not for *P. marinus.*
- This paper demonstrated the potential for 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, via natural selection (e.g., host resistance or tolerance), or to parallel genetic divergence across biogeographic regions (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 one 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

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

individuals infected with hepatitis C virus (HCV) to show how single nucleotide polymorphisms (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 RADseq to characterize host and parasite systems. Bracewell *et al.* (2018) used RADseq to find evidence of co-evolution and cascading speciation among four close interacting species, the western pine 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.* (2018) was able to determine the frequent phenomena of host switching in Panamanian strangler figs and their pollinating fig wasps (Satler *et al.* 2018). A similar occurrence was found in a study by Sweet *et al.* (2020). Sweet *et al.* used whole-genome sequencing and double-digested RAD-sequencing to obtain SNPs to examine co-evolutionary patterns and host-switching accessibility between two 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 *C. virginica* (Ph. Mollusca) along much of the oyster's geographic range from the Gulf of St. Lawrence, Canada to the Gulf of Mexico (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 (USDA 2018). *P. 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 U.S. East Coast and the Gulf of Mexico, 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,

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

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 Gulf of Mexico coast (TX and LA) in oysters collected from all four populations. Reece *et al.* (1997, 2001) genotyped restriction fragment length polymorphisms (RFLP) of *in vitro* cultures of *P. marinus*, noted the presence of diploid or a multiple infections of *P. marinus* in single oysters, and revealed three genetically distinct subdivisions among estuaries of the United States: the northeast Atlantic, southeast Atlantic, and Gulf of Mexico (Reece *et al.*, 1997, 2000). 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 Gulf of Mexico 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 restriction-site associated DNA sequencing (RADseq) to examine single nucleotide polymorphism (SNPs) of *P. marinus* and *C. virginica* within *P. marinus*-infected oysters. Our sampling was focused on infected adults collected at two Gulf of Mexico 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.

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

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 one 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 greater than or equal to 100 ng/uL. If the concentration was greater than 300 ng/uL, then the DNA was diluted in an equal volume of Milli-Q® water before PCR. A concentration of about 100 ng/uL 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 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 ℃ for 30 seconds, annealing 57 ℃ for 30 seconds, and extension at 72 ℃ for 1.5 minutes. 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 minutes and stained with GelRed Nucleic Acid stain in a water bath for 25 minutes. This included a 1KB bp ladder at the first lane of the gel so that the desired 1,000 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.

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

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 uL 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 two 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 two 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), forty-two 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 – 1.01M) was 0.37% of the median number aligned to the *C. virginica* genome (34M; a range of 3.52 – 66.89M). 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 100K to 1M *P. marinus* reads. Another sample identified as 6837-16, from VA had 1.1 M reads,

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

from which we randomly downsampled 500K 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 one 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*,

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

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 analyzed four groups of samples: Alabama (n = 3), Louisiana (n = 9), Virginia Chesapeake Bay (n = 18), 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 analyze genotype likelihoods of *P. marinus* and LEA: SNMF (Frichot and François, 2015) to analyze 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* data set 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 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.

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

RESULTS

Strong genetic divergence between the Gulf of Mexico collections (LA and AL) and Virginia were revealed within the nuclear genomes of both *P. marinus* and *C. virginica* in all analyses. Principal component analysis (PCA; Figure 2), admixture analyses (Figure 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. *C. virginica* genotypes between Chesapeake Bay and seaside Eastern Shore groups were divergent in PCA (Figure 2B), admixture (Figure 3B) and AMOVA (Table 2B). However, *P. marinus* genotypes were not clearly divergent in any of these analyses (Figure 2A, 3A, 4; Table 2A). Regions also did not differ in expected heterozygosity for either *P. marinus* or *C. virginica* (Table 3).

A co-phylogeny plot (Figure 4) provides a view of both the phylogeographic structure and patterns of mixing between host and parasite genotypes. Overall, paco (Hutchinson *et al.*, 2017) 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 Gulf of Mexico versus Atlantic spatial scale. There were two oyster samples collected in GOM (coded LA6946-C and AL7040-6) that had *P. marinus* genotypes that were embedded within the Virginia clade (Figure 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.

DISCUSSION

In our analyses, clear genotypic differentiation was evident between our Atlantic samples and those from the Gulf of Mexico, in both *P. marinus* and *C. virginica* (Figures 2 and 3 and Table 2B). Furthermore, we identified genotypic differentiation within *C. virginica* populations between the Chesapeake Bay and the seaside Eastern Shore in Virginia (Figures 2 and 3 and Table 2B). 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

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

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 (Figure 4) revealed a strong signal of co-phylogeny between the parasite *P. marinus* and its oyster host. A close interaction between these two, documented for nearly three-quarters of a century but potentially of far older origins, has resulted in an ongoing arms race, as these organisms are continually adapting to out compete 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 Gulf of Mexico 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, via pathogen introduction from distant waters via 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

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

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 Gulf of Mexico *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 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's 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.

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

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.* 2018, and 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.

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

Data availability.

Relevant code and datasets are found at https://github.com/esotka/WeatherupPerkinsus
FASTQ files have been uploaded to GenBank (BioSample Accession SAMN39856659-SAMN39856695).

Author's contribution.

All authors conceived and designed the study and edited the final manuscript. Ryan Carnegie and Elizabeth Weatherup collected samples, and Elizabeth Weatherup extracted samples. Elizabeth Weatherup and Erik Sotka 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.

The authors declare there are no conflicts of interest.

Ethical standards.

Not applicable

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

References

Ansari, M. A., Pedergnana, V., L C Ip, C., 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, G. R., Hudson, E., McLauchlan, J., Simmonds, P., Bowden, R., Klenerman, P., Barnes, E. and Spencer, C. C. A. (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. doi: 10.1038/ng.3835.

Bernatchez S, Xuereb A, Laporte M, et al (2019). Seascape genomics of eastern oyster (*Crassostrea virginica*) along the Atlantic coast of Canada. Evolutionary Applications 12:587–609.

Bracewell, R. R., Vanderpool, D., Good, J. M. and Six, D. L. (2018). Cascading speciation among mutualists and antagonists in a tree–beetle–fungi interaction. *Proceedings of the Royal Society B: Biological Sciences* 285, 20180694. doi: 10.1098/rspb.2018.0694.

Burford, M. O., Scarpa, J., Cook, B. J. and Hare, M. P. (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. doi: 10.3354/meps10796.

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. doi: 10.3354/meps139127.

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

Carnegie, R.B., Arzul, I., and Bushek, D. (2016) Managing marine diseases in the context of regional and international commerce: policy essues and emerging concerns. *Philosophical Transactions of the Royal Society B* 371: 20150215. doi:dx.doi.org/10.1098/rstb.2015.0215.

Carnegie, R. B., Ford, S. E., Crockett, R. K., Kingsley-Smith, P. R., Bienlien, L. M., Safi, L. S. L., Whitefleet-Smith, L. A. and Burreson, E. M. (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. doi: 10.1038/s41598-021-92379-6.

Charif, D. and Lobry, J. R. (2007). SeqinR 1.0-2: A Contributed Package to the R Project for Statistical Computing Devoted to Biological Sequences Retrieval and Analysis. In *Structural Approaches to Sequence Evolution: Molecules, Networks, Populations* (ed. Bastolla, U., Porto, M., Roman, H. E., and Vendruscolo, M.), pp. 207–232. Springer, Berlin, Heidelberg doi: 10.1007/978-3-540-35306-5_10.

Choi, Y.-J., Aliota, M. T., Mayhew, G. F., Erickson, S. M. and Christensen, B. M. (2014). Dual RNA-seq of Parasite and Host Reveals Gene Expression Dynamics during Filarial Worm–Mosquito Interactions. *PLOS Neglected Tropical Diseases* 8, e2905. doi: 10.1371/journal.pntd.0002905.

Coen, L. D. and Bishop, M. J. (2015). The ecology, evolution, impacts and management of host-parasite interactions of marine molluscs. *Journal of Invertebrate Pathology* 131, 177–211. doi: 10.1016/j.jip.2015.08.005.

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M. and Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience* 10, giab008. doi: 10.1093/gigascience/giab008.

Day, P.R. (1974). *Genetics of host-parasite interaction*. WH Freeman and Co..

Dexter, E., Fields, P. D. and Ebert, D. (2023). Uncovering the Genomic Basis of Infection Through Co-genomic Sequencing of Hosts and Parasites. *Molecular Biology and Evolution* 40, msad145. doi: 10.1093/molbev/msad145.

Dismukes, W., Braga, M. P., Hembry, D. H., Heath, T. A. and Landis, M. J. (2022). Cophylogenetic Methods to Untangle the Evolutionary History of Ecological Interactions. *Annual Review of Ecology, Evolution, and Systematics* 53, 275–298. doi: 10.1146/annurev-ecolsys-102320-112823.

Ebert, D. and Fields, P. D. (2020). Host-parasite co-evolution and its genomic signature. *Nature Reviews. Genetics* 21, 754–768. doi: 10.1038/s41576-020-0269-1.

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. doi: 10.1111/j.1365-294X.2005.02553.x.

Excoffier, L., Smouse, P. E. and Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131, 479–491. doi: 10.1093/genetics/131.2.479.

Frichot, E. and François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution* 6, 925–929. doi: 10.1111/2041-210X.12382.

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. doi: 10.1093/molbev/msp259.

Hare, M. P. and Avise, J. C. (1996). Molecular genetic analysis of a stepped multilocus cline in the American oyster (*Crassostrea virginica)*. *Evolution; International Journal of Organic Evolution* 50, 2305–2315. doi: 10.1111/j.1558-5646.1996.tb03618.x.

Hutchinson, M. C., Cagua, E. F., Balbuena, J. A., Stouffer, D. B. and Poisot, T. (2017). paco: implementing Procrustean Approach to Cophylogeny in R. *Methods in Ecology and Evolution* 8, 932–940. doi: 10.1111/2041-210X.12736.

Knaus, B. J. and Grünwald, N. J. (2017). vcfr: a package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources* 17, 44–53. doi: 10.1111/1755-0998.12549.

Lees, J. A., Ferwerda, B., Kremer, P. H. C., Wheeler, N. E., Serón, M. V., Croucher, N. J., Gladstone, R. A., Bootsma, H. J., Rots, N. Y., Wijmega-Monsuur, A. J., Sanders, E. A. M., Trzciński, K., Wyllie, A. L., Zwinderman, A. H., van den Berg, L. H., van Rheenen, W., Veldink, J. H., Harboe, Z. B., Lundbo, L. F., de Groot, L. C. P. G. M., van Schoor, N. M., van der Velde, N., Ängquist, L. H., Sørensen, T. I. A., Nohr, E. A., Mentzer, A. J., Mills, T. C., Knight, J. C., du Plessis, M., Nzenze, S., Weiser, J. N., Parkhill, J., Madhi,

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

S., Benfield, T., von Gottberg, A., van der Ende, A., Brouwer, M. C., Barrett, J. C., Bentley, S. D. and van de Beek, D. (2019). Joint sequencing of human and pathogen genomes reveals the genetics of pneumococcal meningitis. *Nature Communications* 10, 2176. doi: 10.1038/s41467-019-09976-3.

Li, H. and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* 26, 589–595. doi: 10.1093/bioinformatics/btp698.

Märkle, H., John, S., Cornille, A., Fields, P. D. and Tellier, A. (2021). Novel genomic approaches to study antagonistic coevolution between hosts and parasites. *Molecular Ecology* 30, 3660–3676. doi: 10.1111/mec.16001.

Meisner, J. and Albrechtsen, A. (2018). Inferring Population Structure and Admixture Proportions in Low-Depth NGS Data. *Genetics* 210, 719–731. doi: 10.1534/genetics.118.301336.

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, D. A., Klinck, J. M., Powell, E. N., Hofmann, E. E., Wilkin, J. and Haidvogel, D. B. (2012). Modeling the dispersal of eastern oyster (*Crassostrea virginica*) larvae in Delaware Bay. *Journal of Marine Research* 70, 381–409. doi: 10.1357/002224012802851940.

Page, R. D. and Charleston, M. A. (1998). Trees within trees: phylogeny and historical associations. *Trends in Ecology & Evolution* 13, 356–359. doi: 10.1016/s0169-5347(98)01438-4.

Paradis, E. (2010). pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics (Oxford, England)* 26, 419–420. doi: 10.1093/bioinformatics/btp696.

R Core Team (2021). R: A language and environment for statistical computing. *R Foundation for Statistical Computing (Vienna, Austria)*. URL https://www.R-project.org/

Reeb, C. A. and Avise, J. C. (1990). A Genetic Discontinuity in a Continuously Distributed Species: Mitochondrial DNA in the American Oyster, *Crassostrea Virginica*. *Genetics* 124, 397–406.

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

Reece, K. S., Bushek, D., Hudson, K. E. and Graves, J. E. (2001). Geographic distribution of *Perkinsus marinus* genetic strains along the Atlantic and Gulf Coasts of the USA. Marine Biology 139: 1047-105

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. doi: 10.1007/s00604-014-1243-4.

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

Ronquist, F. (1997). Phylogenetic approaches in coevolution and biogeography. *Zoologica Scripta* 26, 313–322. doi: 10.1111/j.1463-6409.1997.tb00421.x.

Rose, C. G., Paynter, K. T. and Hare, M. P. (2006). Isolation by Distance in the Eastern Oyster, *Crassostrea virginica*, in Chesapeake Bay. *Journal of Heredity* 97, 158–170. doi: 10.1093/jhered/esj019.

Satler, J. D., Herre, E. A., Jandér, K. C., Eaton, D. A. R., Machado, C. A., Heath, T. A. and Nason, J. D. (2019). Inferring processes of coevolutionary diversification in a community of Panamanian strangler figs and associated pollinating wasps*. *Evolution* 73, 2295–2311. doi: 10.1111/evo.13809.

Shapero, M. H., Leuther, K. K., Nguyen, A., Scott, M. and Jones, K. W. (2001). SNP Genotyping by Multiplexed Solid-Phase Amplification and Fluorescent Minisequencing. *Genome Research* 11, 1926–1934. doi: 10.1101/gr.205001.

Smaal, A. C. and Prins, T. C. (1993). The Uptake of Organic Matter and the Release of Inorganic Nutrients by Bivalve Suspension Feeder Beds. In *Bivalve Filter Feeders* (ed. Dame, R. F.), pp. 271–298. Springer, Berlin, Heidelberg doi: 10.1007/978-3-642-78353-1_8.

Sparks, A. K. (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. doi: 10.4319/lo.1966.11.2.0312.

Stauber, L. A. (1950). The Problem of Physiological Species with Special Reference to Oysters and Oyster Drills. *Ecology* 31, 109–118. doi: 10.2307/1931365.

Sweet, A. D., Wilson, R. E., Sonsthagen, S. A. and Johnson, K. P. (2020). Lousy grouse: Comparing evolutionary patterns in Alaska galliform lice to understand host evolution and host–parasite interactions. *Ecology and Evolution* 10, 8379–8393. doi: 10.1002/ece3.6545

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

Thompson, P. C., Rosenthal, B. M. and Hare, M. P. (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. doi: 10.1016/j.meegid.2011.01.008.

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. doi: 10.2983/035.033.0119.

Thongda, W., Zhao, H., Zhang, D., Jescovitch, L. N., Liu, M., Guo, X., Schrandt, M., Powers, S. P. 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 (New York, N.Y.)* 20, 385–395. doi: 10.1007/s10126-018-9803-y.

Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A. and Minh, B. Q. (2016). W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research* 44, W232-235. doi: 10.1093/nar/gkw256.

Van Valen. L. (1973). A new evolutionary law. *Evolution* 1:1-30.

Varney, R. L., Galindo-Sánchez, C. E., Cruz, P. and Gaffney, P. M. (2009). Population Genetics of the Eastern Oyster *Crassostrea virginica* (Gmelin, 1791) in the Gulf of Mexico. *Journal of Shellfish Research* 28, 855–864. doi: 10.2983/035.028.0415.

Vermeer, K. M. C. A., Dicke, M. and De Jong, P. W. (2011). The potential of a population genomics approach to analyze geographic mosaics of plant--insect coevolution. *Evolutionary Ecology* 25, 977–992. doi: 10.1007/s10682-010-9452-8.

Zimmerman, R. J., Minello, T. J., Baumer, T. and Castiglione, M. C. (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

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

Table 2. Analysis of molecular variance (AMOVA) on *Perkinsus marinus* and *Crassostrea virginica* samples. Overall PhiST (i.e., across four populations) was 0.125 ($p = 0.005$) and 0.165 ($p < 0.001$) for *P. marinus* and *C. virginica*, respectively. Pairwise-PhiST and values are shown below and above diagonals, respectively.

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	

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

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

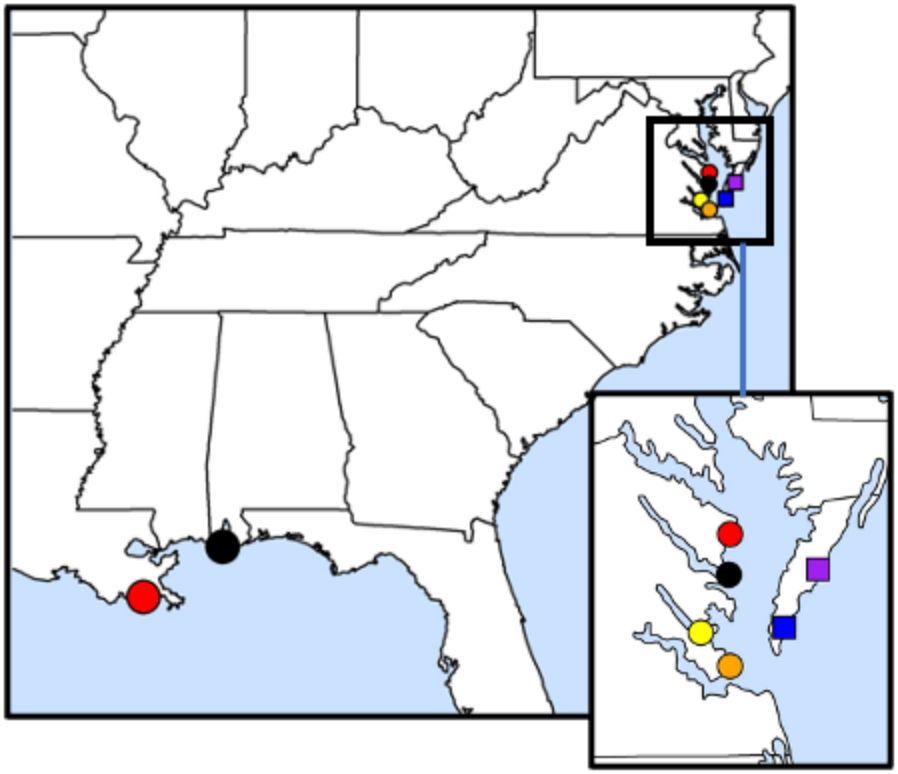

Fig 1. Map of geographic location of each site collection for the region of Virginia (circle = Chesapeake Bay; squares = Eastern Shore; Red = Fleet Point (n=2); Yellow = VIMS Beach (n=11); Orange = Wreck Shoal (n= 1); Black = Broad Creek (n=1); Purple = Wachapreague (n=1); Blue = Oyster (n=6)) and two regions in the Gulf of Mexico (Red = Louisiana (n=5) and Black = Alabama (n=3)).

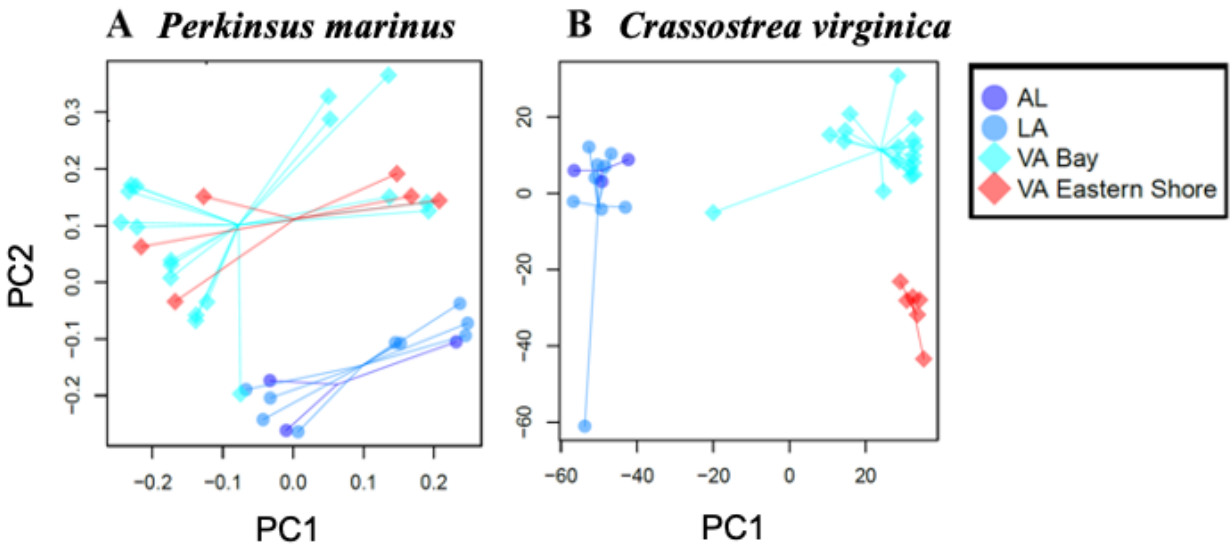

Fig. 2. Principal components analyses of A) *Perkinsus marinus* (393 loci, n = 36) from PCangsd and B) *Crassostrea virginica* (52100 SNPs, n = 35) from prcomp.

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

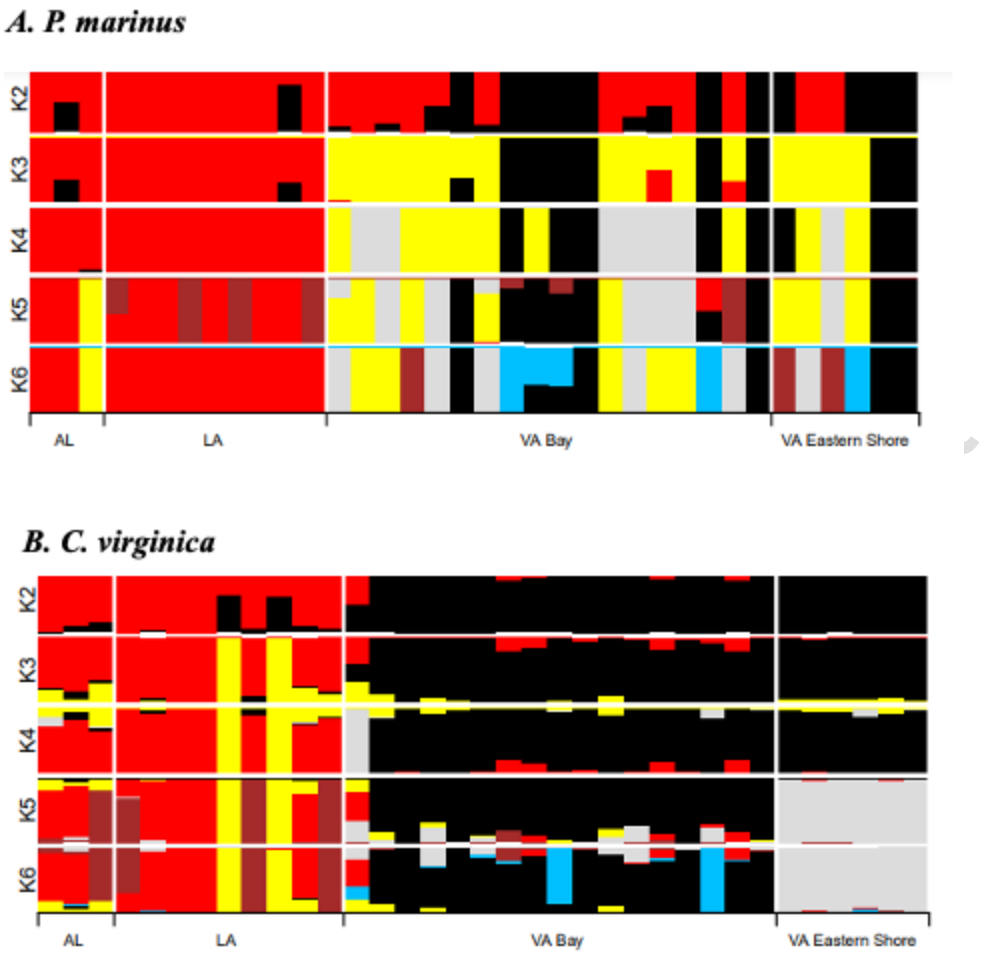

Fig. 3. Admixture analysis of A) *Perkinsus marinus* (393 loci, n = 36) using genotype likelihoods in NGSadmix and B) *Crassostrea virginica* (52100 SNPs, n = 35) using genotypes in SNMF.

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

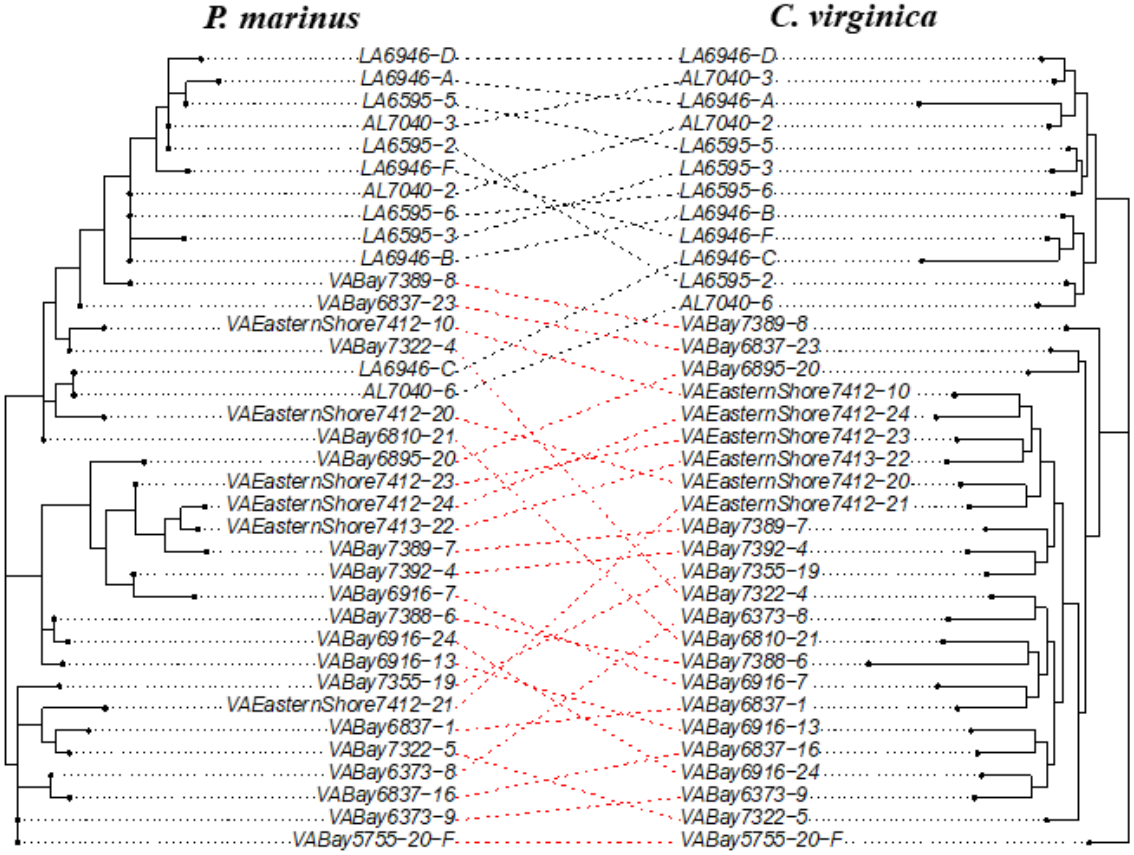

Fig. 4. A maximum-likelihood co-phylogeny of the parasite *Perkinsus marinus* (392 bp) and host *Crassostrea virginica* (52052 bp). All nodes have 100% consensus support for *C. virginica* while all nodes have <50% support for *P. marinus* (1000 bootstrap replicates for both). Black and red dashed lines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

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