Little evidence for genetic variation associated with susceptibility to sea star wasting syndrome in the keystone species, *Pisaster ochraceus*

Andrea Burton[1], Sarah Gravem[1], and Felipe Barreto[1]

[1]Oregon State University

February 22, 2024

Abstract

The keystone species, *Pisaster ochraceus*, suffered mass mortalities along the northeast Pacific Ocean from Sea Star Wasting Syndrome (SSWS) outbreaks in 2013-2016. SSWS causation is still debated, leading to concerns as to whether outbreaks will continue to impact this species. Considering the apparent link between ocean temperature and SSWS, the future of this species and intertidal communities remains uncertain. Surveys of co-occurring *P. ochraceus* along the central Oregon coast in 2016 allowed us to address whether variation in disease status showed genetic variation that may be associated with differences in susceptibility to SSWS. We performed restriction site-associated DNA sequencing (2bRAD-seq) to genotype ˜72,000 of single nucleotide polymorphism (SNP) loci across apparently normal and wasting sea stars. Locus-specific analyses of differentiation (Fst) between disease-status groups revealed no signal of genetic differences separating the two groups. Using a multivariate approach, we observed weak separation between the groups, but identified 18 SNP loci showing highest discriminatory power between the groups and scanned the genome annotation for linked genes. A total of 34 protein-coding genes were found to be located within 15 kb (measured by linkage disequilibrium decay) of at least one of the 18 SNPs, and 30 of these genes had homologies to annotated protein databases. Our results suggest that likelihood of developing SSWS symptoms does not have a strong genetic basis. The few genomic regions highlighted had only modest levels of differentiation, but the genes associated with these regions may form the basis for functional studies aiming to understand disease progression.

The keystone species, *Pisaster ochraceus*, suffered mass mortalities along the northeast Pacific Ocean from Sea Star Wasting Syndrome (SSWS) outbreaks in 2013-2016. SSWS causation is still debated, leading to concerns as to whether outbreaks will continue to impact this species. Considering the apparent link between ocean temperature and SSWS, the future of this species and intertidal communities remains uncertain. Surveys of co-occurring *P. ochraceus* along the central Oregon coast in 2016 allowed us to address whether variation in disease status showed genetic variation that may be associated with differences in susceptibility to SSWS. We performed restriction site-associated DNA sequencing (2bRAD-seq) to genotype ˜72,000 of single nucleotide polymorphism (SNP) loci across apparently normal and wasting sea stars. Locus-specific analyses of differentiation (Fst) between disease-status groups revealed no signal of genetic differences separating the two groups. Using a multivariate approach, we observed weak separation between the groups, but identified 18 SNP loci showing highest discriminatory power between the groups and scanned the genome annotation for linked genes. A total of 34 protein-coding genes were found to be located within 15 kb (measured by linkage disequilibrium decay) of at least one of the 18 SNPs, and 30 of these genes had homologies to annotated protein databases. Our results suggest that likelihood of developing SSWS symptoms does not have a strong genetic basis. The few genomic regions highlighted had only modest levels of differentiation, but the genes associated with these regions may form the basis for functional studies aiming to understand disease progression.

Little evidence for genetic variation associated with susceptibility to sea star wasting syndrome in the keystone species, *Pisaster ochraceus*

Andrea Burton[1], Sarah Gravem[1], and Felipe Barreto[1]

[1]Department of Integrative Biology, Oregon State University, Corvallis, Oregon, U.S.A

*To whom correspondence should be addressed: Andrea Burton, Integrative Biology, Cordley Hall, 2701 SW Campus Way, Corvallis, OR 97331, USA,

Email: burtoand@oregonstate.edu

Abstract

The keystone species, *Pisaster ochraceus*, suffered mass mortalities along the northeast Pacific Ocean from Sea Star Wasting Syndrome (SSWS) outbreaks in 2013-2016. SSWS causation is still debated, leading to concerns as to whether outbreaks will continue to impact this species. Considering the apparent link between ocean temperature and SSWS, the future of this species and intertidal communities remains uncertain. Surveys of co-occurring *P. ochraceus* along the central Oregon coast in 2016 allowed us to address whether variation in disease status showed genetic variation that may be associated with differences in susceptibility to SSWS. We performed restriction site-associated DNA sequencing (2bRAD-seq) to genotype ˜72,000 of single nucleotide polymorphism (SNP) loci across apparently normal and wasting sea stars. Locus-specific analyses of differentiation (Fst) between disease-status groups revealed no signal of genetic differences separating the two groups. Using a multivariate approach, we observed weak separation between the groups, but identified 18 SNP loci showing highest discriminatory power between the groups and scanned the genome annotation for linked genes. A total of 34 protein-coding genes were found to be located within 15 kb (measured by linkage disequilibrium decay) of at least one of the 18 SNPs, and 30 of these genes had homologies to annotated protein databases. Our results suggest that likelihood of developing SSWS symptoms does not have a strong genetic basis. The few genomic regions highlighted had only modest levels of differentiation, but the genes associated with these regions may form the basis for functional studies aiming to understand disease progression.

Keywords

Marine disease, sea star wasting syndrome (SSWS), allele frequency, keystone species

Introduction

Little evidence for genetic variation associated with susceptibility to sea star wasting syndrome in the keystone species, *Pisaster ochraceus*, Rising sea water temperatures, due to climate change, are becoming increasingly stressful to marine ecosystems. As a result, marine diseases have become more prevalent in the last few decades (Harvell et al., 2004, 2002). Disease outbreaks can have detrimental downstream effects on marine species due to changes in community structure and age distribution (Behringer, Silliman, & Lafferty, 2018; Burge et al., 2014). Many marine taxa have suffered intense population declines as a result of the increasing prevalence of diseases (Harvell et al. 2004, Burge et al. 2013). These declines can result in reduced variation due to population bottlenecks and genetic drift, and possibly from strong directional selection associated with tolerance or resistance to disease (Nei, Maruyama, & Chakrabort, 1975; Zenger, Richardson, & Vachot-Griffin, 2003).

The Sea Star Wasting Syndrome (SSWS) epidemic event that began in 2013 is believed to be the largest marine wildlife disease on record (Gravem et al., 2021; Harvell et al., 2019). SSWS affected over 20 species of sea stars from Baja California, Mexico to the Gulf of Alaska, United States (Hewson et al., 2018, 2014), and severely reduced population sizes of several sea star species (Gravem et al., 2021; Harvell et al., 2019; Hewson et al., 2014; Menge, Cerny-Chipman, Johnson, & Sullivan, 2016; Miner et al., 2018; Montecino-Latorre et al., 2016). Similar SSWS symptoms have been observed in British Columbia in 2008 (Bates, Hilton, & Harley, 2009), along the US east coast (Bucci et al., 2017), the South Pacific (Pratchett, 1999; Zann, Brodie, & Vuki, 1990), Australia and Yellow Sea (Hewson et al., 2019). While the viral candidate sea

star-associated densovirus (SSaDV) has been debunked (Jackson et al., 2020), other hypotheses of causative or exacerbating agents remains unknown, with hypotheses including pathogen(s) (Lloyd & Pespeni, 2018), inconsistent etiology stress responses between location, species, and environment (Hewson et al., 2018), microbial dysbiosis (Lloyd & Pespeni, 2018), and microbial-driven depletion of oxygen at the animal-water interface (Aquino et al., 2021). There is also mixed evidence for whether anomalously warm waters linked to global warming initiated the outbreak (Eisenlord et al., 2016; Menge, Cerny-Chipman, Johnson, Sullivan, et al., 2016b; Miner et al., 2018; Tracy, Weil, & Harvell, 2020) (Aalto et al., 2020). Regardless, it is clear that the disease is exacerbated in warmer conditions (Bates et al., 2009; Eckert, Engle, & Kushner, 1999; Eisenlord et al., 2016; Kohl, McClure, & Miner, 2016), and that severe population reductions occurred in warmer southern regions (Gravem et al., 2021; Harvell et al., 2019; Miner et al., 2018). The interplay between climate change and disease is a growing threat to wildlife species, especially when it causes rapid and extreme populations decline. What is still unclear is whether tolerance or resistance to some of these diseases has a genetic bases that may allow populations to adapt if outbreaks continue to occur.

The keystone species *Pisaster ochraceus* was severely affected by SSWS over much of its range. In Oregon, their populations declined by 50-94%. Because *P. ochraceus* aids in maintaining fast-growing Mytilus californianus populations from overgrowing intertidal zones (Paine, 1966, 1969, 1974), declines in their populations have resulted in trophic cascades and regime shifts in intertidal regions (Burt et al., 2018; Miner et al., 2018; Schultz, Cloutier, & Cote, 2016). Loss of *P. ochraceus* due to SSWS could have detrimental impacts on coastal ecosystems. It is likely that SSWS has exerted strong selection on *P. ochraceus* populations. Recent genetic studies on this and other affected sea star species are suggestive of a genetic component to variation in SSWS susceptibility. Individuals of *P. ochraceus* with SSWS symptoms showed elevated expression levels in genes associated with immune response and tissue remodeling (Fuess et al., 2015; Gudenkauf & Hewson, 2015; Ruiz-Ramos, Schiebelhut, Hoff, Wares, & Dawson, 2020). In addition, Schiebelhut et al. (2018) observed allele frequencies shifts before and after peak SSWS outbreaks in California populations. More specifically, they detected changes in restriction site-associated DNA sequencing (RAD-seq) haplotype frequencies between pre-SSWS adults and post-SSWS adults, as well as between pre- SSWS adults and recruits in the populations after SSWS. These changes occurred in few loci, but were consistent across independent geographic samples (Schiebelhut, Puritz, & Dawson, 2018). However, because Schiebelhut et al. (2018) genotyped only apparently normal individuals (asymptomatic), it is still unclear whether the allele shifts were caused by the disease itself or by other co-occurring factors.

Here, we build upon their work by investigating genomic differences between wasting individuals (i.e., presenting with SSWS) and grossly, or apparently normal individuals from the same localities during an outbreak of SSWS. We examine genetic variation in 200 *P. ochraceus* individuals collected in central Oregon in 2016, two years following the initial spring 2014 SSWS outbreak in Oregon (Menge et al. 2016). At this time, both apparently normal and wasting sea stars were common at each of the six Oregon sites sampled. We reasoned that, by being found on the same transects as wasting individuals and hence likely exposed to similar conditions, sea stars found to be apparently normal may carry genetic variants associated with resistance or tolerance to SSWS. By combining field surveys of natural disease prevalence with high-throughput single nucleotide polymorphism (SNP) genotyping, we assess the contribution of sea star genetic variation to SSWS occurrence. Our dataset is the result of a unique opportunity to compare apparently normal and wasting individuals from the same time and place during the SSWS epidemic.

Materials and Methods

All sea stars were collected between April and October 2016 across six sites in central Oregon. These included two northern sites ˜0.7 km apart near Depoe Bay, at Fogarty Creek on 20 Oct 2016 (44.8386, -124.0588) and Boiler Bay on 23 May 2016 (44.8303, -124.0608). The two central sites sampled were Smelt Sands (44.3212, -124.1081, on 16 Aug 2016) and Yachats Beach (44.3114, -124.1086, on 8 May 2016); theses were 55 km southward from Boiler Bay, and 0.2 km apart from one another. The two southern sites were 7.4 km southward on Cape Perpetua and were 4.7 km from one another; these were Strawberry Hill on 1 Jul 2016 (44.2492, -124.1154) and Tokatee Klootchman on 24 Apr 2016 (44.2037, -124.1170).

Animals were collected by hand at low tide. Arm length were recorded (center to longest arm) for each animal, and only adults (with >3cm arm length) were scored for disease status and had tissue sampled. We recorded disease symptoms as per Menge et al. (2016); these included, in order of severity: lesions, arm loss, twisted arms, deflated (loss of turgor pressure), and disintegrating/dying. Animals were considered healthy if none of these symptoms nor any injuries existed (e.g. regrowing arms). Tube feet (˜5-10) were collected from each animal using scrubbed and sterilized forceps, then stored in 1.5mL microcentrifuge tubes containing 1 mL of 95% ethanol. All samples were stored on ice and then at –20°C until ready for DNA isolation.

For genotyping, we included only individuals with the highest symptomatic scores in order to avoid misdiagnosis, as well to include only individuals that were truly not resistant to SSWS (i.e. not in the process of recovery after mild disease). In total, 82 individuals were included in the symptomatic group and 112 in the asymptomatic group, but these proportions ranged across sites (Table 1). DNA was extracted using the E.Z.N.A Tissue DNA kit (Omega Biotek, Norcross, GA) and quantified using a fluorescence method (Quant-iT dsDNA Assay Kit, ThermoFisher, Waltham, MA). We used the 2bRAD protocol for genotyping SNPs (Wang et al. 2012), following the original published protocol, but using the enzyme AlfI. We also used adaptors with "NN" overhangs to target 100% of restriction sites. Multiplexed individuals were pooled at approximately equimolar amounts (after quantification via quantitative PCR) and sequenced across five lanes of an Illumina HiSeq 3000 as 50-bp single reads.

Adaptors were trimmed and low-quality reads were filtered (<30 phred) using publicly available scripts (https://github.com/Eli-Meyer/2brad_utilities/). Cleaned reads were mapped to the reference *P. ochraceus* genome (NCBI; GCA_010994315.1) using SHRiMP (Rumble et al., 2009), reporting the top three maximum hits per read. We used Stacks v.1.35 (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013; Catchen, Amores, Hohenlohe, Cresko, & Postlethwait, 2011) to call genotypes using default parameters, with samples coded by disease status group (symptomatic/asymptomatic). Additional filtering parameters used included: selection of a single (first) SNP per stack, removal of loci that were genotyped in only one group, the minimum minor allele frequency (MAF) set to 0.025, and a minimum minor allele count set to four. Finally, the percentage of individuals needed to process a locus were adjusted for each population so that a minimum of eight samples were genotyped in a group. In VCFtools (Danecek et al., 2011), retained only biallelic SNPs and loci with minimum depth of 6.

Material and Methods

Field tissue sampling

Field surveys aimed at quantifying prevalence of SSWS were conducted between April and July 2016 in the low intertidal zone at five sites in central Oregon: Fogarty Creek (44.8386, - 124.0588), Boiler Bay (44.8303, -124.0608), Yachats Beach (44.3114, -124.1086), Strawberry Hill (44.2492, -124.1154), and Tokatee Klootchman (44.2037, -124.1170). These surveys were conducted using 5m x 2m belt transects (5-10 transects per site). Animals were collected by hand at low tide. Arm length was recorded (center to longest arm) for each animal, and only adults (with >3cm arm length) were scored for disease status (Menge, Cerny-Chipman, Johnson, Sullivan, et al., 2016a). We recorded visual disease symptoms based on the six-level ranking protocol, as per Menge et al. (2016); these included, in order of severity: twisting arms (1), deflated (2), lesions (3), lost arms (4), losing grip on rocks (5), and disintegrating or "melting" (6). Animals were considered apparently normal if none of these symptoms existed.

We returned to each of these sites to collect tissue for genetic analysis (Table S1). We also collected tissue at a sixth site (Smelt Sands, 44.3212, -124.1081, on 16 Aug 2016), but we did not conduct transect surveys there. From each adult individual, tube feet (˜5-10) were collected using scrubbed and sterilized forceps, then stored in 1.5-mL microcentrifuge tubes containing 1 mL of 95% ethanol. All samples were stored on ice and then at –20?C until ready for DNA isolation. In total, we collected tissue from 410 sea stars, 92 of which were wasting.

Library preparation and sequencing

For genotyping, we included only individuals with the highest wasting scores from each site, which ranged from ranks 3 to 6 (Table S1). In total, 82 individuals were included in the wasting group and 112 in the apparently normal group. DNA was extracted using the E.Z.N.A Tissue DNA kit (Omega Biotek, Norcross, GA) and quantified using a fluorescence method (Quant-iT dsDNA Assay Kit, ThermoFisher, Waltham, MA). We used the 2bRAD protocol for genotyping SNPs (Wang et al. 2012), following the original published protocol, but using the enzyme AlfI. We also used adaptors with "NN" overhangs to target 100% of restriction sites. Multiplexed individuals were pooled at approximately equimolar amounts (after quantification via quantitative PCR) and sequenced across five lanes of an Illumina HiSeq 3000 as 50-bp single reads, at the Center for Genome Research and Biocomputing at Oregon State University.

Data filtering

Adaptors were trimmed and low-quality reads were filtered (<30 phred) using publicly available scripts (https://github.com/Eli-Meyer/2brad_utilities/). Because of the cleavage pattern of the AlfI enzyme, 2bRAD DNA inserts are 34-36 bp in length. Therefore, after adaptor and quality trimming, we filtered out any reads that were shorter than 34 bp to reduce chances of mismapping. Cleaned reads were mapped to the reference *P. ochraceus* genome (NCBI accession GCA_010994315.1) using SHRiMP (Rumble et al., 2009), reporting the top three maximum hits per read. We used Stacks v.1.35 (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013; Catchen, Amores, Hohenlohe, Cresko, & Postlethwait, 2011) to call genotypes using default parameters, with samples coded by disease status group (wasting/apparently normal). Additional filtering parameters used included: selection of a single (first) SNP per stack, removal of loci that were genotyped in only one group, the minimum minor allele frequency (MAF) set to 0.025, a minimum minor allele count set to four, and only loci represented in at least 50% of samples were retained. Using VCFtools (Danecek et al., 2011), we retained only biallelic loci and only genotypes with a minimum coverage of 6 reads. Finally, we used PLINK (Purcell et al., 2007) to remove individuals that were missing more than 50% of loci. This filtering pipeline retained a dataset with 133 individuals (74 normal and 59 wasting) and 71,784 SNP loci, which we will refer to as the 'full dataset'.

Population structure

Before comparing genotypic variation between sea stars varying in disease status, we assessed whether significant population genetic structure exists among the sampled sites. We used the Bayesian clustering approach in STRUCTURE (Pritchard, Stephens, & Donnelly, 2000) with the dataset of 133 individuals, but we further filtered it to reduce computational burden. For this, we filtered loci that were genotyped in at least 90% of individuals, using VCFtools, retaining 4,626 loci. STRUCTURE runs included 100,000 burn-in and 100,000 MCMC sampling replicates, assuming admixture and with sampling locality used as prior information. We ran three values of K to assess whether clustering occurring at the specific site (K = 6) or at the subregion-level (K = 3) were more likely than a large panmictic population (K = 1). To confirm that parameters converged, each value of K was run three separate times and likelihoods and clustering patterns compared across runs. This clustering analysis suggested genetic variation was not structured among sampling sites (Figure S1). We therefore considered our samples as coming from a single population in the subsequent analyses.

Analyses of genetic variation

We used two types of approaches for estimating differentiation between disease status groups. We estimated Weir & Cockerham's Fst (Weir & Cockerham, 1984) using GPAT++ (Shapiro et al., 2013), with significance levels adjusted using the false discovery rate (FDR) in R. To test for Fst outliers that may be indicative of selection, we used BayeScan v2.1 (Foll & Gaggiotti, 2008) with the following settings: prior odds of the neutral model of 10, burn-in of 50,000 replicates, a thinning interval of 10, 20 pilot runs for 5,000 iterations, and recorded output of 5,000 iterations. Significance for Bayescan Fst outliers was assessed at a q-value false-discovery rate (FDR) of 0.01.

We also used a multivariate approach to test for differentiation between grossly normal and wasting sea stars. A discriminant analyses of principal components (DAPC) (Jombart, Devillard, & Balloux, 2010) was used,

implemented in the R package 'adegenet' (Jombart, 2015) and identified outlier loci based on their loadings associated with the discriminant function separating the two groups.

Regions linked to discriminant loci

To identify functional candidates that may be associated with disease status, we scanned for protein-coding genes linked to outlier SNPs. We first determined the appropriate genomic window size for scanning around each SNP by estimating linkage disequilibrium (LD) between pairs of SNPs. The r2 was calculated using VCFtools between pairs of SNPs within 100,000-bp windows. Average r2 was calculated in bins of 100-bp increments, and were plotted against physical distance. This plot showed that LD decays rapidly up to 15 kb, then continues to decrease but at lower rates (Figure S2). We hence scanned a 30-kb window centered at each outlier SNP (15 kb on either side) by overlaying the protein-coding genome annotation from Ruiz-Ramos et al. (2020) onto the genome assembly, using the Integrative Genomics Viewer (Robinson et al., 2011).

Results

Field observations

A total of 3,670 *P. ochraceus* were counted in belt transects across five sites, with incidence of SSWS ranging from 5.2% to 10.0% (Table 1). In three of the sites (Boiler Bay, Yachats Bay, and Strawberry Hill), all surveyed sea stars showed either symptoms of lesions or arm loss, while at Tokatee Klootchman, 25% showed the more advanced symptom of grip loss. Fogarty Creek harbored sea stars with the most advanced stage of SSWS; all wasting individuals surveyed were disintegrating (Table 1). Metadata associated with the 410 individuals from which we collected tissue sample can be found on Table S1.

Sequencing and genotyping

Illumina sequencing yielded ˜1.6 billion reads. After removing low quality reads 1.4 billion reads remained, with an average of 6.9 million reads per sample. A total of 259,407 RAD stacks passed sample and population filters, surveying a total of 9,526,221 bases (˜2.3% of the 401.9 Mb genome) and with mean per-sample coverage of 38.1x. As mentioned above, 71,784 RAD stacks contained at least one polymorphic site, and a single random site was retained per stack in the full dataset (Table S2). Moreover, given our estimated LD block of 15 kb, this full SNP set allowed for an average of 2.5 SNPs sampled per block. This level of coverage suggests our data set may provide sufficient power to detect genomic regions associate with phenotypic differences, if these exist and have a strong genetic component (Lowry et al., 2017a, 2017b).

Analyses of genomic variation

Genomic differentiation between grossly normal and wasting sea stars was very low based on Fst estimates. Across the final SNP dataset (71,784), Weir & Cockerham's Fst had a median value of 0.00314, and nearly 45% of loci had Fst = 0 (Table S2). Moreover, while 362 loci had moderate Fst values ([?] 0.1), no locus showed significant differentiation after FDR adjustments; the lowest adjusted P-value was 0.172 (Figure 1). Outlier tests with Fst using Bayescan also showed no evidence of selection in any locus in our dataset (Figure S3).

DAPC analyses showed modest separation between apparently normal and wasting groups (Figure 2A). Based on loading values from the DAPC, we identified 18 SNPs across 10 chromosomes contributing most to the differentiation between the two groups (Figure 2B, Table S3). Allele frequency differences across these loci ranged from 0.014 - 0.253 (Table S3). Using a 30-kb window centered at each of these SNP positions, we detected 34 protein-coding genes predicted by the genome annotation from Ruiz-Ramos et al. (2020). BLAST of these protein sequences against the Uniprot/Swissprot databased revealed that 30 of them have predicted products with known functional annotation (Table S4). Chromosomes 3 and 8 harbored the most genes linked to these SNPs (seven genes each); in chromosome 8, one SNP was linked to three genes, and the others to two each (Figure 3).

We assessed whether the 18 SNP loci identified as outliers from our DAPC analyses overlapped with haplotypes from Schiebelhut et al. (2018) that were the most discriminatory between post- and pre-SSWS adults in their samples from California. For this, we compared the 30-kb ranges encompassing our SNP outliers to the haplotype coordinates from Schiebelhut et al. (2018). We detected no overlap in the two sets of outliers, with the shortest distance detected as ˜257 kb (Table S5).

Discussion

The recent outbreaks of SSWS at multiple coastal sites caused severe population declines in several sea star species. Mitigation techniques for addressing outbreaks when causative agent(s) is unknown should run parallel with studies attempting to determine the cause (Groner et al., 2016). With the rising sea water temperatures resulting in the higher prevalence of marine diseases (Harvell et al., 2002; Tracy, Pielmeier, Yoshioka, Heron, & Harvell, 2019), we are likely to see similar scenarios of mass mortality outbreaks impacting marine species more frequently and having little time to address management or conservation plans. Assessing the potential for natural population resilience is a critical step in predicting the long-term fate of affected species and of the communities they in turn influence. For example, selectively rearing disease-resistant oysters in hatcheries has been a useful tool in avoiding disease outbreaks that are decimating wild populations (Agnew et al., 2020; Degremont, Garcia, & Allen, 2015). While many marine species are not amenable for selective breeding, examining genomic variation in natural populations can address whether these species have the genetic makeup for adaptation to marine diseases on their own.

In this study, we took advantage of the co-occurrence of wasting and apparently normal individuals of *P. ochraceus* in central Oregon to scan for genomic regions that potentially predict individual SSWS status. After genotyping nearly 72,000 SNP loci across 133 individuals, we found no strong patterns of differentiation between wasting and apparently normal individuals. Loci with elevated Fst were not clearly concentrated as peaks in any genomic region, and no single locus showed statistically significant level of allele frequency differences. Using a multivariate approach as a complement to the locus-specific Fst analyses, 18 SNP loci stood out as contributing to genomic differentiation between the two groups of individuals based on disease status. Overall, we argue that a genetic basis for SSWS resilience in *P. ochraceus* is likely weak, but we identified a list of genomic regions and functional candidates that may serve as basis for studies of gene expression, physiology, or comparative genomics during future SSWS outbreaks.

While the proximate cause(s) of SSWS at the individual level are still unknown, recent experimental studies are consistent with a pathogen agent. For example, individuals that were wasting in the laboratory showed physiological and gene expression response suggestive of innate immunity, cytokine-like systems, and tissue remodeling (Fuess et al., 2015; Gudenkauf & Hewson, 2015; Ruiz-Ramos et al., 2020). Our findings showed no evidence for strong genetic component to SSWS tolerance or resistance, but the weakly-associated loci we identified may have small but cumulative effects, which is expected for a polygenic trait. This trait may hence require much higher powered studies for detecting associated loci with more precision (Gagnaire & Gaggiotti, 2016).

Schielbelhut et al. (2018) detected allelic shifts in grossly normal *P. ochraceus* adults and juveniles before and after the SSWS outbreak in California. They found three loci putatively under selection and reported on 100 discriminatory haplotypes between time periods. Interestingly, the 18 SNP loci we detected as most discriminatory in our samples did not occur within 30 kb of those reported by Schielbelhut et al. (2018), and most were between 250 kb and 2 Mb apart. Lack of overlap in these genomic regions is perhaps not surprising given the multitude of differences between the studies, such as year of sampling, the health status of sea stars, and geographic location. For instance, SSWS is known to be associated with spikes in sea water temperature (Bates et al., 2009; Eckert et al., 1999; Eisenlord et al., 2016; Kohl et al., 2016), and daily deviations from annual sea water temperatures in Oregon were more prevalent in 2013/2014 than in California (Miner et al., 2018). Such relevant environmental differences between California and Oregon may hence also cause different selective pressures. Moreover, the reduced-representation nature of RADseq, and the use of different restriction enzymes suggest that a lack of overlap between studies' outlier markers is not indicative of a lack of biological relevance.

The scarcity of SSWS-associated loci isolated is also possibly a result of reduced coverage from the inherent nature of RADseq methods (Lowry et al., 2017a). While RADseq is an efficient and cost-effect method for producing thousands of SNPs along the entire genome, these markers remain sparse. Despite the limitations, many RADseq studies have found loci attributing to adaptive selection when coverage is adequate (Epstein et al., 2016; Lowry et al., 2017a; McKinney, Larson, Seeb, & Seeb, 2017). Marker density aimed at detecting phenotype-genotype associations is recommended to be high and relative to LD in the target species (Lowry et al., 2017a). Based on this metric, our RADseq effort in this study adequately covers the full genome, with on average 2.5 SNPs found per every 15-kb linkage block. Therefore, we argue that our results are not due to low marker density, but perhaps may be improved by genotyping a higher number of individuals.

Our study joins that of Schielbelhut et al. (2018) and Ruiz-Ramos et al. (2020) in assessing genomic variation in the keystone species *P. ochraceus* and highlighting the importance of understanding the causes and responses to devastating SSWS outbreaks. While current patterns remain obscure, the accumulation of putative functional genomic regions will serve as invaluable resources for continued field and laboratory studies. In addition to physiological and transcriptomic experiments, we suggest the need for a concerted effort to sample large numbers of wasting and apparently unaffected individuals across several geographic regions, and ideally using low-coverage whole-genome sequencing for substantially increased power (Lou, Jacobs, Wilder, & Therkildsen, 2021).

Acknowledgments

We thank Kris Bauer, Laurel Field, Skylar Peyen, Bruce Menge, Brittany Poirson, Amanda Coração, and Jonathan Robinson for help with sample collections and DNA extractions. Thanks to M. Dasenko for assistance with Illumina sequencing. We would also like to thank Lauren Schiebelhut for access to the genome assembly. We would also like to thank Eli Meyer who provided support on 2bRAD preparation and processing. This work was supported by Oregon State University startup funds to F.S.B.

References

Aalto, E. A., Lafferty, K. D., Sokolow, S. H., Grewelle, R. E., Ben-Horin, T., Boch, C. A., ... De Leo, G. A. (2020). Models with environmental drivers offer a plausible mechanism for the rapid spread of infectious disease outbreaks in marine organisms. *Scientific Reports* , *10* (1), 5975. https://doi.org/10.1038/s41598-020-62118-4

Agnew, M. V., Friedman, C. S., Langdon, C., Divilov, K., Schoolfield, B., Morga, B., ... Burge, C. A. (2020). Differential mortality and high viral load in naive pacific oyster families exposed to OsHV-1 suggests tolerance rather than resistance to infection.*Pathogens* , *9* (12), 1057. https://doi.org/10.3390/pathogens9121057

Aquino, C. A., Besemer, R. M., DeRito, C. M., Kocian, J., Porter, I. R., Raimondi, P. T., ... Hewson, I. (2021). Evidence that microorganisms at the animal-water interface drive sea star wasting disease. *Frontiers in Microbiology* , *11* . https://doi.org/10.3389/fmicb.2020.610009

Bao, Z., & Eddy, S. R. (2002). Automated de novo identification of repeat sequence families in sequenced genomes. *Genome Research* ,*12* (8), 1269–1276. https://doi.org/10.1101/gr.88502

Behringer, D. C., Silliman, B. R., & Lafferty, K. D. (Eds.). (2018).*Marine Disease Ecology* . Londong: Oxford University Press.

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* , *27* (2), 573–580. https://doi.org/10.1093/nar/27.2.573

Burge, C. A., Mark Eakin, C., Friedman, C. S., Froelich, B., Hershberger, P. K., Hofmann, E. E., ... Harvell, C. D. (2014). Climate Change Influences on Marine Infectious Diseases: Implications for Management and Society. *Annual Review of Marine Science* ,*6* (1), 249–277. https://doi.org/10.1146/annurev-marine-010213-135029

Burt, J. M., Tinker, M. T., Okamoto, D. K., Demes, K. W., Holmes, K., & Salomon, A. K. (2018). Sudden

collapse of a mesopredator reveals its complementary role in mediating rocky reef regime shifts.*Proceedings of the Royal Society B: Biological Sciences* ,*285* (1883), 20180553. https://doi.org/10.1098/rspb.2018.0553

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks : Building and genotyping loci de novo from short-read sequences. *G3* , *1* (3), 171–182. https://doi.org/10.1534/g3.111.000240

Catchen, J. M., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics.*Molecular Ecology* , *22* (11), 3124–3140. https://doi.org/10.1111/mec.12354

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... Group, 1000 Genomes Project Analysis. (2011). The variant call format and VCFtools. *Bioinformatics* ,*27* (15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Degremont, L., Garcia, C., & Allen, S. K. (2015). Genetic improvement for disease resistance in oysters: A review. *Journal of Invertebrate Pathology* , *131* , 226–241. https://doi.org/10.1016/j.jip.2015.05.010

Eisenlord, M. E., Groner, M. L., Yoshioka, R. M., Elliott, J., Maynard, J., Fradkin, S., ... Harvell, D. C. (2016). Ochre star mortality during the 2014 wasting disease epizootic: Role of population size structure and temperature. *Philosophical Transactions of the Royal Society B: Biological Sciences* , *371* (1689). https://doi.org/10.1089/rstb.2015.0212

Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* , *180* (2), 977–993. https://doi.org/10.1534/genetics.108.092221

Fuess, L. E., Eisenlord, M. E., Closek, C. J., Tracy, A. M., Mauntz, R., Gignoux-Wolfsohn, S., ... Roberts, S. B. (2015). Up in arms: Immune and nervous system response to sea star wasting disease. *PLoS ONE* , *10* (7), e0133053. https://doi.org/10.1371/journal.pone.0133053

Gagnaire, P.-A., & Gaggiotti, O. E. (2016). Detecting polygenic selection in marine populations by combining population genomics and quantitative genetics approaches. *Current Zoology* , *62* (6), 603–616. https://doi.org/10.1093/cz/zow088

Gravem, S. A., Heady, W. N., Saccomanno, V. R., Alvastad, K. F., Gehman, A. L. M., Frierson, T. N., & Hamilton, S. L. (2021). Sunflower Sea Star (Pycnopodia helianthoides). *IUCN Red List of Threatened Species* .

Groner, M. L., Maynard, J., Breyta, R., Carnegie, R. B., Dobson, A., Friedman, C. S., ... Harvell, C. D. (2016). Managing marine disease emergencies in an era of rapid change. *Philosophical Transactions of the Royal Society B: Biological Sciences* , *371* (1689), 20150364. https://doi.org/10.1098/rstb.2015.0364

Gudenkauf, B. M., & Hewson, I. (2015). Metatranscriptomic analysis of Pycnopodia helianthoides (Asteroidea) affected by sea star wasting disease. *PLoS ONE* , *10* (5). https://doi.org/10.1371/journal.pone.0128150

Harvell, D. C., Aronson, R., Baron, N., Connell, J., Dobson, A., Ellner, S., ... Ward, J. (2004). The rising tide of ocean diseases: unsolved problems and research priorities. *Frontiers in Ecology and the Environment* , *2* (7), 375–382.

Harvell, D. C., Mitchell, C. E., Ward, J. R., Altizer, S., Dobson, A. P., Ostfel, R. S., & Samuel, M. D. (2002). Climate warming and disease risks for terrestrial and marine biota. *Science* ,*296* (5576), 2158–2162.

Harvell, D. C., Montecino-Latorre, D., Caldwell, J. M., Burt, J. M., Bosley, K., Keller, A., ... Gaydos, J. K. (2019). Disease epidemic and a marine heat wave are associated with the continental-scale collapse of a pivotal predator (Pycnopodia helianthoides). *Science Advances* , *5* (1), eaau7042. https://doi.org/10.1126/sciadv.aau7042

Hewson, I., Bistolas, K. S. I., Quijano Carde, E. M., Button, J. B., Foster, P. J., Flanzenbaum, J. M., ... Lewis, C. K. (2018). Investigating the complex association between viral ecology, environment, and northeast Pacific Sea Star Wasting. *Frontiers in Marine Science* , *5* (MAR). https://doi.org/10.3389/fmars.2018.00077

Hewson, I., Button, J. B., Gudenkauf, B. M., Miner, B., Newton, A. L., Gaydos, J. K., ... Harvell, D. C. (2014). Densovirus associated with sea-star wasting disease and mass mortality. *Proceedings of the National Academy of Sciences of the United States of America* ,*111* (48), 17278–17283. https://doi.org/10.1073/pnas.1416625111

Holt, C., & Yandell, M. (2011). MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects.*BMC Bioinformatics* , *12* (1), 491. https://doi.org/10.1186/1471-2105-12-491

Jombart, T. (2015). An introduction to adegenet 2.0.0. Retrieved from http://adegenet.r-forge.r-project.org/files/tutorial-basics.pdf

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* , *11* (1), 94. https://doi.org/10.1186/1471-2156-11-94

Kohl, W. T., McClure, T. I., & Miner, B. G. (2016). Decreased Temperature Facilitates Short-Term Sea Star Wasting Disease Survival in the Keystone Intertidal Sea Star *Pisaster ochraceus*. *PLoS ONE* ,*11* (4), e0153670. https://doi.org/10.1371/journal.pone.0153670

Korf, I. (2004). Gene finding in novel genomes. *BMC Bioinformatics* , *5* , 59. https://doi.org/10.1186/1471-2105-5-59

Lloyd, M. M., & Pespeni, M. H. (2018). Microbiome shifts with onset and progression of Sea Star Wasting Disease revealed through time course sampling. *Scientific Reports* , *8* (1), 16476. https://doi.org/10.1038/s41598-018-34697-w

MCCALLUM, H., KURIS, A., HARVELL, C., LAFFERTY, K., SMITH, G., & PORTER, J. (2004). Does terrestrial epidemiology apply to marine systems? *Trends in Ecology & Evolution* , *19* (11), 585–591. https://doi.org/10.1016/j.tree.2004.08.009

Menge, B. A., Cerny-Chipman, E. B., Johnson, A., & Sullivan, J. (2016). Correction: Sea Star Wasting Disease in the Keystone Predator *Pisaster ochraceus* in Oregon: Insights into Differential Population Impacts, Recovery, Predation Rate,. *PloS One* , *11* (6), e0157302. Retrieved from http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0153994

Menge, B. A., Cerny-Chipman, E. B., Johnson, A., Sullivan, J., Gravem, S., & Chan, F. (2016). Sea Star Wasting Disease in the Keystone Predator *Pisaster ochraceus* in Oregon: Insights into Differential Population Impacts, Recovery, Predation Rate, and Temperature Effects from Long-Term Research. *PLoS ONE* , *11* (6), e0157302. https://doi.org/10.1371/journal.pone.0153994

Miner, C. M., Burnaford, J. L., Ambrose, R. F., Antrim, L., Bohlmann, H., Blanchette, C. A., ... Raimondi, P. T. (2018). Large-scale impacts of sea star wasting disease (SSWD) on intertidal sea stars and implications for recovery. *PLoS ONE* , *13* (3), e0192870. https://doi.org/10.1371/journal.pone.0192870

Montecino-Latorre, D., Eisenlord, M. E., Turner, M., Yoshioka, R., Harvell, C. D., Pattengill-Semmens, C. V., ... Gaydos, J. K. (2016). Devastating transboundary impacts of sea star wasting disease on subtidal asteroids. *PLoS ONE* , *11* (10), e0163190. https://doi.org/10.1371/journal.pone.0163190

Mydlarz, L. D., Jones, L. E., & Harvell, C. D. (2006). Innate immunity, environmental drivers, and disease ecology of marine and freshwater invertebrates. *Annual Review of Ecology, Evolution, and Systematics* , *37* (1), 251–288. https://doi.org/10.1146/annurev.ecolsys.37.091305.110103

Nei, M., Maruyama, T., & Chakrabort, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution* , *29* , 1–10.

Paine, R. T. (1966). Food web complexity and species diversity. *The American Naturalist* , *100* (910), 91–75.

Paine, R. T. (1969). A note on trophic complexity and community stability. *The American Naturalist* , *103* , 91–93.

Paine, R. T. (1974). Intertidal community structure. *Oecologia* ,*15* , 93–120.

Price, A. L., Jones, N. C., & Pevzner, P. A. (2005). De novo identification of repeat families in large genomes. *Bioinformatics* , *21* , 351–358. https://doi.org/10.1093/bioinformatics/bti1018

Ruiz-Ramos, D. V., Schiebelhut, L. M., Hoff, K. J., Wares, J. P., & Dawson, M. N. (2020). An initial comparative genomic autopsy of wasting disease in sea stars. *Molecular Ecology* , *29* (6), 1087–1102. https://doi.org/10.1111/mec.15386

Rumble, S. M., Lacroute, P., Dalca, A. V., Fiume, M., Sidow, A., & Brudno, M. (2009). SHRiMP: Accurate mapping of short color-space reads. *PLoS Computational Biology* , *5* (5), e1000386. https://doi.org/10.1371/journal.pcbi.1000386

Schiebelhut, L. M., Puritz, J. B, & Dawson, M. N. (2018. Decimation by sea star wasting disease and rapid genetic change in a kystone species, *Pisaster ochraceus. Proceedings of the National Academy of Sciences of the United States of America* , *115* , 7069-7074. https://doi.org/10.1073/pnas.1800285115

Schultz, J. A., Cloutier, R. N., & Cote, I. M. (2016). Evidence for a trophic cascade on rocky reefs following sea star mass mortality in British Columbia. *PeerJ* , *4* , e1980. https://doi.org/10.7717/peerj.1980

Shapiro, M. D., Kronenberg, Z., Li, C., Domyan, E. T., Pan, H., Campbell, M., ... Wang, J. (2013). Genomic diversity and evolution of the head crest in the rock pigeon. *Science* , *339* (6123), 1063–1067. https://doi.org/10.1126/science.1230422

Slater, G. S. C., & Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* , *6* (31). https://doi.org/10.1186/1471-2105-6-31

Smit, A., & Hubley, R. (2014). RepeatModeler Open-1.0.

Stanke, M., & Morgenstern, B. (2005). AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Research* , *33* , 465–467. https://doi.org/10.1093/nar/gki458

Tracy, A. M., Pielmeier, M. L., Yoshioka, R. M., Heron, S. F., & Harvell, D. C. (2019). Increases and decreases in marine disease reports in an era of global change. *Proceedings of the Royal Society B: Biological Sciences* , *286* (1912), 20191718. https://doi.org/10.1098/rspb.2019.1718

Tracy, A. M., Weil, E., & Harvell, D. C. (2020). Warming and pollutants interact to modulate octocoral immunity and shape disease outcomes. *Ecological Applications* , *30* (2), e02024. https://doi.org/10.1002/eap.2024

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-Statistics for the Analysis of Population Structure. *Evolution* , *38* (6), 1358–1370. https://doi.org/10.2307/2408641

Zenger, K. R., Richardson, B. J., & Vachot-Griffin, A.-M. (2003). A rapid population expansion retains genetic diversity within European rabbits in Australia. *Molecular Ecology* , *12* (3), 789–794. https://doi.org/10.1046/j.1365-294X.2003.01759.x

Data Accessibility

Illumina sequence reads are deposited in the NCBI Sequence Read Archive (SRA), in accessions SRR13611638 - SRR13611837.

Author Contributions

S.G and F.B conducted field sampling. A.B. undertook sample preparation with help from F.B. A.B. and F.B. performed processing and analysis of data. F.B. completed the genome annotation. A.B. wrote manuscript with critical feedback and editing by F.B. and S.G.

Hosted file

`image1.emf` available at https://authorea.com/users/409999/articles/519510-little-evidence-for-genetic-variation-associated-with-susceptibility-to-sea-star-wasting-syndrome-in-the-keystone-species-pisaster-ochraceus

Figure 1. Discriminant analysis of principal components for symptomatic versus asymptomatic sea stars within the (A) north, (B) middle, and (C) south populations.

Hosted file

`image2.emf` available at https://authorea.com/users/409999/articles/519510-little-evidence-for-genetic-variation-associated-with-susceptibility-to-sea-star-wasting-syndrome-in-the-keystone-species-pisaster-ochraceus

Figure 2. Fst values for SNP loci with outlier DAPC loadings. Plotted are loci that were outliers in the DAPC analysis within each population in comparisons between symptomatic and asymptomatic sea stars. Vertical bands show the ten regions (100 kbp) within which outlier loci from two or more populations were clustered.

Table 1. Number of adult *P. ochraceus* sample collections from the three geographic populations. Individuals were noted as symptomatic, asymptomatic or recovering from SSWS. A subset of symptomatic and asymptomatic individuals were used for sequencing (in parenthesis).

Population	Symptomatic	Asymptomatic	Recovering	Total
North	14 (14)	38 (38)	17	69 (52)
Middle	18 (18)	77 (43)	10	105 (61)
South	50 (50)	37 (37)	29	116 (87)
Total	82 (82)	152 (118)	56	290 (200)

Table 2. Allele frequencies of three SNP loci detected as outliers in BayeScan Fst tests between symptomatic and asymptomatic *P. ochraceus* . Shown are the frequencies of the reference allele in each of the geographic regions. Bold-face values are those frequencies that were significantly different at FDR = 0.01.

Chromosome	Position		Allele Frequency	Allele Frequency	Allele Frequency	Number genotyped
			North	Middle	South	South
15	6,870,781	Symptomatic	0.833	0.666	**1**	**1**
		Asymptomatic	0.675	0.789	**0.333**	**0.333**
11	7,983,511	Symptomatic	**0.291**	0.8	0.647	0.647
		Asymptomatic	**0.896**	0.783	0.773	0.773
22	4,132,115	Symptomatic	**0.923**	0.708	0.518	0.518
		Asymptomatic	**0.29**	0.5	0.712	0.712