Big Fishery, Big Data, and Little Crabs: Using Genomic Methods to Examine the Seasonal Recruitment Patterns of Early Life Stage Dungeness Crab (*Cancer magister*) in the California Current Ecosystem

Elizabeth M. J. Lee† and Kathleen G. O’Malley†

State Fisheries Genomics Lab, Coastal Oregon Marine Experiment Station, Department of Fisheries and Wildlife, Hatfield Marine Science Center, Oregon State University, Newport, OR, United States

The California Current Ecosystem (CCE) is a dynamic marine ecosystem from which many socioeconomically important fisheries species are harvested. Here, a genotyping-by-sequencing (GBS) approach was used to examine genomic variation in an early life stage (megalopae) of the Dungeness crab (*Cancer magister*), which constitutes the most valuable single-species commercial fishery in the CCE. Variation in abundance and timing of megalopae recruitment has been extensively studied for over two decades in Coos Bay, Oregon, United States. Within the CCE, documented timing of Dungeness crab life history events indicates that coastal megalopae recruitment is expected to occur April through July; however, long-term studies in Coos Bay have observed late-season recruitment from August to October. Based on variation at 1,913 presumably neutral loci, evidence was found for weak, yet significant differentiation (\(F_{ST}\) estimate = 0.0011) between the 2014 expected-season recruits (\(n = 47\)) and late-season recruits (\(n = 47\)) collected in Coos Bay. However, two putatively adaptive loci with a high \(F_{ST}\) estimate (0.2036) between expected-season and late-season recruits were identified. These findings support the hypothesis that expected-season and late-season megalopae recruiting to Coos Bay within the same year may have originated from different locations or from different breeding groups. Understanding marine species connectivity between ecosystems is important when considering how future changes in ocean conditions may impact fishery harvests.

Keywords: population genomics, population connectivity, genotyping-by-sequencing, ocean conditions, California Current Ecosystem, larval transport, larval dispersal, Dungeness crab

INTRODUCTION

The California Current Ecosystem (CCE) is a dynamic marine ecosystem that spans from Vancouver Island, British Columbia, Canada to Baja California, Mexico (Fautin et al., 2010). This large marine ecosystem exhibits spatial and temporal variations in ocean conditions, such as sea surface temperature, sea surface height, timing of spring transition, alongshore winds, and
The Dungeness crab (Cancer magister, Dana, 1852, following the naming convention currently recognized by the Integrated Taxonomic Information System) is considered the most valuable single-species commercial fishery within the CCE, regularly earning the highest annual coastwide ex-vessel value (Rasmuson, 2013; California Department of Fish and Wildlife, 2018; Oregon Department of Fish and Wildlife, 2018; Washington Department of Fish and Wildlife, 2018). Dungeness crab landings within the CCE have fluctuated annually, by an order of magnitude, since the establishment of the commercial fishery in the early 1900s (Rasmuson, 2013). At present, managers regulate the CCE Dungeness crab commercial fishery using the “3-S strategy,” size, sex, and season, which limits harvest to only male crab of a certain carapace width during a specified season (Fisher and Velasquez, 2008; Rasmuson, 2013; California Department of Fish and Wildlife, 2018; Oregon Department of Fish and Wildlife, 2018). Although methods for aging adult Dungeness crab currently do not exist, catch models approximate that, on average, the CCE crab reach legal size for commercial harvest at age four (Botsford, 1984; Johnson et al., 1986). Therefore, under the current management strategy, it is estimated that greater than 90% of all age four male crabs are harvested annually; indicating that the coastwide CCE Dungeness crab commercial fishery landings (Washington, Oregon, and California, United States) are a reasonable proxy for the size of the age 4 year class (Hackett et al., 2003). Furthermore, the CCE harvest variability has often been attributed to larval success of each year class, when the Dungeness crab are most vulnerable to ocean conditions (reviewed in Botsford et al., 1989; reviewed in Botsford and Lawrence, 2002; reviewed in Rasmuson, 2013).

Although the Dungeness crab is characterized as benthic at harvest, its life history begins with a long pelagic larval duration (PLD) of 3–4 months. Reproducing female crab brood their eggs until hatching, at which time the first pelagic larval stage (zoea) are moved offshore and dispersed along the coast by two main current systems within the CCE. First, zoea are transported northward by the Davidson Current (pre-spring transition; northern flowing coastal winds) and then southward by the California Current (post-spring transition; southern flowing coastal winds) (Wild and Tasto, 1983; Hackett et al., 2013). During the PLD, zoea develop through five stages before metamorphosing into a final megalopae stage (Wild and Tasto, 1983). Recruitment of the megalopae from offshore back to the nearshore occurs during coastal upwelling, when the megalopae migrate toward the coast with internal tides and then successfully settle in the nearshore environment (Jamieson and Phillips, 1988; Johnson and Shanks, 2002; Rasmuson, 2013). Migration is thought to be limited after the recruiting megalopae settle (i.e., less than 20 km a year) (Hildenbrand et al., 2011). Based on growth and survival rates within the local nearshore habitat, male Dungeness crab reproduce at least once before reaching harvestable size within the CCE fishery (McKelvey et al., 1980; Hackett et al., 2003).

Dungeness crab inhabit a large geographic area, which extends northward from the CCE to the Aleutian Islands, Alaska, United States, encompassing the Salish Sea Ecosystem (SSE) and the Gulf of Alaska (GOA) ecosystem (Wild and Tasto, 1983; Sherman and McGovern, 2012). Within and among these ecosystems, the timing of life history events differs along a latitudinal gradient (Figure 1) (reviewed in Rasmuson, 2013). For example, within the CCE, SSE, and GOA, mating occurs March–June, April–September, and June–July, respectively (Strathmann, 1987; Fisher, 2006). Likewise, subsequent egg brooding and larval hatch timing differs latitudinally. Within the southern portion of the CCE (coastal California), larval hatching occurs December–February and within the northern portion of the CCE (coastal Oregon and Washington), larval hatching occurs January–March. Furthermore, hatching occurs February–May within the SSE and much later in the year (i.e., June–July) within the GOA (Strathmann, 1987; Fisher, 2006).

Documented studies of early life history of the Dungeness crab are spatially limited across the species’ large geographic distribution, but the recruitment timing and abundance of the Dungeness crab megalopae has been closely studied over the past two decades in Coos Bay, Oregon (Shanks and Roenger, 2007; Shanks et al., 2010; Shanks, 2013). Throughout the species range, regional nearshore recruitment and settlement of megalopae generally occurs 3–4 months after hatching; thus, matching the documented 3–4 month PLD of the species (Rasmuson, 2013). However, research in Coos Bay has found that megalopae recruitment continues later into the year than expected with recruits settling through August and into late October in some years (Shanks and Roenger, 2007; Shanks et al., 2010; Shanks, 2013). These late-season megalopae recruits do not fit the documented life history timing for this species.

Ocean conditions play an important role in larval transport, dispersal, survival, and overall population connectivity (Cowen and Sponaugle, 2009). The previous studies of Dungeness crab megalopae recruitment in Coos Bay have demonstrated that the Pacific Decadal Oscillation (PDO), physical spring transition, and upwelling events, are correlated with annual megalopae recruitment abundances of the Dungeness crab. For instance, the abundance of megalopae recruits in Coos Bay is higher during years of a negative PDO, earlier spring transition, and stronger upwelling (Shanks and Roenger, 2007; Shanks et al., 2010; Shanks, 2013). Accordingly, annual megalopae recruitment abundances vary annually by orders of magnitude in Coos Bay. Interestingly, Shanks et al. (2010) and Shanks (2013) also found a density dependent correlation between the annual megalopae recruitment abundances in Coos Bay and the total coastwide CCE Dungeness crab commercial landings 4 years later, when the megalopae recruitment cohort is expected to reach harvestable size (Hackett et al., 2003).

It has been hypothesized that the Dungeness crab megalopae recruiting to Coos Bay later in the year (late-season: August–September) than expected (expected-season: April–July), are from the northern portion of the species distribution such as the SSE or the GOA (Shanks, 2013). This hypothesis is based...
on several observations of the pelagic larval stage. First, higher abundances of Coos Bay late-season megalopae recruits are observed in negative PDO years, when the southern flowing California Current is stronger (Shanks, 2013). During negative PDO years, more water is diverted from the North Pacific Gyre to the California Current than to the GOA Current (Keister et al., 2011; Shanks, 2013); therefore, greater southern transport of larvae could occur during stronger negative PDO years. Second, the August-September timing of the late-season megalopae recruitment in Coos Bay aligns with the timing of megalopae recruitment in the two northern ecosystems of the species range, the SSE and the GOA. Third, morphologically, the Coos Bay late-season recruits are smaller than the expected-season recruits (Shanks et al., 2010), and this is consistent with studies of juvenile Dungeness crab in coastal Washington (CCE) and Puget Sound (SSE) where three intra-annual recruitment events have been observed based on differences in recruitment timing and physical size of megalopae recruits (Dinnel et al., 1993). The researchers found that the recruitment events that occurred later in the year (from August to September) was composed of smaller recruits, and concluded these smaller recruits had originated from within the Puget Sound (SSE) (Dinnel et al., 1993).

Previous genetics studies analyzing 10 neutral microsatellite loci have found evidence for genetic differentiation among the Dungeness crab from different ecosystems along the west coast of North America. For example, O’Malley et al. (2017) found evidence for strong genetic differentiation ($F_{ST} > 0.16$) between benthic stage Dungeness crab in coastal Oregon (CCE) and in British Columbia (samples from both the southern GOA and the northern SSE). A second study utilizing the same methods, but different geographic sampling locations, found evidence for weak, yet significant genetic differentiation ($F_{ST}$ range = 0.002–0.004) between the benthic stage Dungeness CCE and in Puget Sound (southern SSE) (Jackson and O’Malley, 2017). The degree of population connectivity among benthic stage Dungeness crab has also been shown to vary inter-annually. Jackson et al. (2017) found evidence for genetic variation among CCE benthic adults in 2012 but not in 2014 ($F_{ST}$ range = 0.000–0.015). Temporal variation in genetic connectivity can be a result of chaotic genetic patchiness (Johnson and Black, 1982); however, the authors hypothesized that the recruiting year classes experienced different CCE ocean conditions (i.e., spring transition and upwelling) resulting in an isolation by distance (IBD) signal in 1 year but not in the other year.

Here, we present the first study of Dungeness crab megalopae using a genotyping-by-sequencing (GBS) approach. Specifically, we examined megalopae recruiting in 2014 to Coos Bay, Oregon, a geographic location where megalopae recruitment abundances have been extensively studied by others (Shanks and Roegner, 2007; Shanks et al., 2010; Shanks, 2013). Moreover, the findings were placed in the context of the 2014 ocean conditions (i.e., positive PDO, later spring transition, and below average upwelling; Table 1) as well as the CCE commercial fishery landings during the 2018 fishery season. The primary objective of this study was to test for genetic differentiation between expected-season and late-season megalopae recruits, using both presumably neutral and putatively adaptive loci to inform inferences about the origin of late-season recruits. We hypothesized that the expected-season and late-season recruits would be genetically differentiated based on variation at presumably neutral and putatively adaptive loci because the
late-season recruits originated from outside the CCE, likely the SSE or the GOA.

**MATERIALS AND METHODS**

**Larval Sample Collection**

Following the methods of Shanks et al. (2010), a light trap was deployed in Coos Bay, Oregon, United States (43.34°N, 124.33°W) (Figure 2) in 2014 to collect and enumerate daily Dungeness crab megalopae recruits. The deployment time period spanned the expected-season (April–July) and late-season (August–September) recruitment for Dungeness crab megalopae, and daily counts of megalopae were recorded (Figure 3). During large recruitment events, 50 megalopae were subsampled from 1 day and preserved in 95% ethanol for genomic analyses. The carapace length and width were measured for each sampled megalopae. Despite a positive PDO, later spring transition, and below average upwelling, recruits were observed in both the expected-season and the late-season.

**Library Preparation and Sequencing**

To investigate intra-annual differentiation among Dungeness crab recruits, 47 individual megalopae were sampled during the peak of the expected-season (May) and 47 individual megalopae were sampled during the peak of the late-season (September) (Figure 3) (unpublished data Shanks, 2018). Each individual megalopae was homogenized using a TissueLyser II (Qiagen, Hilden, Germany) and genomic DNA was extracted using the Qiagen DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany).

A GBS approach was used to sequence a set of loci across the Dungeness crab genome. Libraries consisting of 96 samples (94 unique megalopae individuals and two replicates) were constructed following the methods of Elshire et al. (2011), with the modification of double digesting the samples with the high-fidelity restriction enzymes SbfI andMspI for 1 h at 37°C. Adapters with 96 unique barcodes (5–10 base pairs in length) were ligated to the 96 samples before multiplexing, to allow for individual sample identification in downstream analyses. Size selection was not performed on the library before polymerase chain reaction (PCR) amplification with Illumina sequencing primers. A Qiagen QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) was used to purify the library and an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, United States) was used to assess the quality of the library. The library was sequenced on an Illumina HiSeq 3000 (Illumina, San Diego, CA, United States) flow cell lane with 150 base pair (bp) paired-end sequencing chemistry. Raw, unfiltered sequencing reads were recorded in a FASTQ format file (Cock et al., 2009).

**Quality Filtering and Genotyping**

The raw sequencing reads were filtered before conducting population genomic analyses. The program FASTQC (v.0.11.3, Andrews, 2010) was used to assess the quality of Illumina raw reads and CUTADAPT (v.1.5, Martin, 2011) was used to trim any Illumina adapters from the raw sequences. The STACKS (v.2.2, Catchen et al., 2011, 2013) process_radtags program was used to further filter and demultiplex the paired-end reads (-P) using the inline unique barcodes (-inline_null). The process_radtags parameters were set to remove reads with an uncalled base (-c), correct barcodes and restriction enzyme sites that were one mismatch different from a true barcode or restriction site sequence (-r), and remove reads that dropped below an average quality score of 20 (-q 20) within a sliding window of 15% of the read length (-w 0.15).

Genotyping the individual megalopae using the filtered short-read sequences was accomplished using the STACKS (v.2.2, Catchen et al., 2011, 2013) software pipeline (Rochette and Catchen, 2017). The STACKS denovo_map.pl program was used to execute each of the programs in the STACKS pipeline: ustacks, cstacks, sstacks, tsv2bam, gstacks, and populations. Parameters for each program in the pipeline were optimized for the Dungeness
crab megalopae samples following the recommendations of Mastretta-Yanes et al. (2015) and Paris et al. (2017). The usstacks program (\texttt{-M 2, -m 3, -N 4}) was used to \textit{de novo} assemble matching reads within an individual to form a set of putative loci using a maximum likelihood framework. The cstacks program (\texttt{-n 1}) was used to compile a catalog of all loci within the samples. The sstacks program was used to match all individuals to the compiled catalog. The tsv2bam program was used to transpose the data. Finally, the gstacks program was used to assemble paired-end reads into contigs using the catalog of loci and then call single nucleotide polymorphisms (SNPs) within each locus.

Population Genomic Analyses
The \textit{STACKS} (v.2.2, Catchen et al., 2011, 2013) \textit{populations} program was used to compare loci between expected-season and late-season megalopae recruits (Rochette and Catchen, 2017; Paris et al., 2017). The \textit{populations} program identified loci found in both expected-season and late-season samples (\texttt{-p 2}) and found in at least 70\% of individuals in both expected-season and late-season samples (\texttt{-r 0.7}). Additionally, loci with minor allele frequencies greater than 0.05 (\texttt{-\textit{min\_maf} 0.05}) were retained. Only the first SNP in each locus was used for downstream analyses to maintain independence of loci (\texttt{-write\_single\_snp}). Results of the \textit{populations} module were output as a Variant Call Format (VCF) file (Danecek et al., 2011). Putative Paralogous Sequence Variants (PSVs) were identified within the expected-season and late-season groups using the program \textit{VCFTOOLS} (v.0.1.13, Wiggin\textit{ton} et al., 2005; Danecek et al., 2011), and a false discovery rate (FDR) correction was applied because of multiple testing (Storey et al., 2004). The R package \textit{HIERFSTAT} (v.3.5.0, Goudet, 2005; R Core Team, 2018) was used to calculate allelic richness (\textit{A}R), observed heterozygosity (\textit{H}O), expected heterozygosity (\textit{H}E), and inbreeding coefficients (\textit{F}IS) within expected-season and late-season sampling groups (Nei and Chesser, 1983; Nei, 1987; El Mousadik and Petit, 1996).

Identifying Putatively Adaptive Loci
Two methods were used to identify putatively adaptive loci considered to be under selective pressure within expected-season and late-season samples. Loci with high \textit{F}ST values (outlier loci) were identified using the program \textit{BAYESCAN} (v.2.1, Foll and Gaggiotti, 2008) with a prior of 100 and a false discovery rate threshold (\textit{q\_value}) of 0.05. With the program \textit{OUTFLANK} (v0.2, Whitlock and Lotterhos, 2015), loci outside an inferred neutral distribution of \textit{F}ST values (outlier loci) were identified with a false discovery threshold (\textit{q\_value}) of 0.01. Putatively adaptive loci sequences were compared against the National Center for Biotechnology Information (NCBI) database using BLASTN (Nucleotide Basic Local Alignment Search Tool) (Altschul et al., 1990) and the \textit{Somewhat similar sequences} program selection.

Testing for Differentiation
To test for differentiation between the expected-season and late-season megalopae recruits, pairwise \textit{F}ST estimates were computed for presumably neutral and putatively adaptive loci.
The R packages adegenet (v. 3.5.0, Jombart, 2008; Jombart and Ahmed, 2011; R Core Team, 2018) and stamp (v. 3.5.0, Pemberton et al., 2013; R Core Team, 2018) were used to calculate pairwise $F_{ST}$ estimate using 10,000 bootstraps across presumably neutral loci, putatively adaptive loci, and all loci to determine 95% confidence intervals and $p$-values (Weir and Cockerham, 1984). The R package adegenet was also used for a Discriminate Analysis of Principle Components (DAPC) across all loci (v. 3.5.0, Jombart et al., 2010; R Core Team, 2018).

RESULTS

Larval Sample Collection

In 2014, 197,242 megalopae were caught in a light trap deployed in Coos Bay, Oregon between April 1 and September 30 (Figure 3) (unpublished data Shanks, 2018). Megalopae were caught on the first day the light trap was deployed, April 10. The trap was checked daily through September 23, and the last day megalopae were caught was September 17. During this time period, daily megalopae catch abundances fluctuated from 0 to 47,600. A total of 170,833 megalopae recruits (87%) were caught within the expected-season (April–July) and a total of 26,562 megalopae recruits (13%) were caught within the late-season (August–September). The largest peak in abundance was observed in May during expected-season (maximum daily count of 47,600 on May 7) and the second largest peak in abundance was observed in September during the late-season (maximum daily count of 21,700 on September 3). Forty-seven megalopae from each peak were used for the genomic analyses. The mean carapace length of 21,700 on September 3). Forty-seven megalopae from

![Mean carapace length of megalopae recruits sampled for genomic analysis from Coos Bay in May (expected-season; $n = 47$) and September (late-season; $n = 47$).](image)

FIGURE 4 | Mean carapace length of megalopae recruits sampled for genomic analysis from Coos Bay in May (expected-season; $n = 47$) and September (late-season; $n = 47$).

Library Sequencing, Filtering, and Genotyping

Genomic DNA was successfully extracted from all 47 expected-season and all 47 late-season megalopae and GBS libraries were constructed and sequenced. The 402,561,387 paired reads (805,122,774 total reads) that resulted from sequencing were evaluated for read quality with FASTQC. Of these paired reads, 146,482,358 (36.39%) were removed due to the presence of an Illumina adapter. Read quality was further assessed using the stacks process_radtags program. In total, 33,559,227 paired reads without barcodes, 40,492,029 paired reads without a restriction enzyme cut site, and 14,828,129 paired reads with low quality were removed from the dataset. Demultiplexing the reads with the sample barcodes revealed that the 94 individual megalopae were adequately represented with a mean of 322,785 ($± 80,967$ SE) reads per individual.

Similar sequences for each of the 94 individuals were aligned into putative loci with the stacks usstacks program and combined into a catalog of 354,735 putative loci using the cstacks program. The catalog of consensus sequences was compared against each individual’s stacks with stacks, tsv2bam was used to transform the data, and finally, gstacks built contigs from the paired-end data and called 5,811 SNPs.

Population Genomic Analyses

The stacks populations program identified 2,216 putative polymorphic loci between expected-season and late-season recruits. After removing loci identified as PSVs, 1,915 loci remained. There was no evidence of linkage disequilibrium between pairs of loci within expected-season and late-season sampling groups. Of the 1,915 loci, 333 loci within the expected-season group and 322 loci within the late-season group showed significant deviation from Hardy–Weinberg Proportion after correction for multiple testing. Removal of the loci out of HWP did not affect the results of downstream genetic differentiation analyses. However, these loci were retained in the final set of 1,915 loci because we were interested in examining intra-annual genetic structure and not the genetic structure within each seasonal group.

The set of 1,915 loci were used for population genetic analyses. Allelic richness for expected-season recruits and late-season recruits was 1.998 and 1.997, respectively (Table 2). Observed heterozygosity was 0.2299 for expected-season recruits and 0.2321 for late-season recruits (Table 2). Expected heterozygosity was higher for both the expected-season recruits (0.2726) and for the late-season recruits (0.2729) (Table 2). The inbreeding

| Population genetic measure | Expected-season | Late-season |
|----------------------------|----------------|------------|
| Allelic richness ($\hat{A}_r$) | 1.9998 | 1.9973 |
| Observed heterozygosity ($H_e$) | 0.2299 | 0.2321 |
| Expected heterozygosity ($\hat{H}_e$) | 0.2726 | 0.2729 |
| Inbreeding coefficient ($F_{IS}$) | 0.1455 | 0.1414 |

![TABLE 2 | Population genetic measurements for the 2014 expected-season and late-season megalopae recruits in Coos Bay, Oregon.](image)
coefficient ($F_{IS}$) was slightly higher for expected-season recruits (0.1455) than late-season recruits (0.1414) (Table 2), but the values were not significantly different ($p > 0.05$; Kruskal–Wallis rank sum test). $F_{IS}$ was significantly different from zero for both expected-season and late-season ($p < 0.01$; one-sample t-test) indicating that individuals in each group are more related than expected under a model of random mating.

**Identification of Putatively Adaptive Loci**

Using the program BAYESCAN, two putatively adaptive loci were identified. The program OUTFLANK also identified these two loci as well as nine additional loci. The two loci identified by both programs were categorized as putatively adaptive loci for downstream analyses. Using BLASTN, one of the putatively adaptive loci matched a hypothetical protein mRNA sequence of the Asian mud crab (Scylla paramamosain) in the NCBI database with an e-value of 5e-28 (GenBank Sequence ID: HM217907.1).

**Differentiation Between Expected-Season and Late-Season Recruits**

The expected-season megalopae recruits were significantly differentiated from the late-season megalopae recruits based on variation at 1,913 presumably neutral loci ($F_{ST} = 0.0011$; $p = 0.0262$; Table 3). The degree of differentiation between the two recruiting groups was much larger based on variation at the two putatively adaptive loci ($F_{ST} = 0.2036$; $p < 0.001$; Table 3). The DAPC using both the presumably neutral and putatively adaptive loci (1,915 loci) indicated that two groups (expected-season and late-season) could be differentiated using five or more principle components (Figure 5). Accordingly, the expected-season and late-season megalopae were also significantly differentiated based on variation at both the presumably neutral and putatively adaptive loci ($F_{ST} = 0.0013$; $p = 0.0109$; Table 3).

**DISCUSSION**

**Differentiation Between Dungeness Crab Megalopae Recruits**

We found evidence for weak genetic differentiation, based on variation at 1,913 presumably neutral loci, between the 2014 expected-season (May) and late-season (September) Dungeness crab megalopae recruits in Coos Bay, Oregon, an estuary within the CCE. We hypothesized that expected-season and late-season recruits would be different because previous studies (1) suggest that the CCE late-season recruits originate from northern ecosystems (GOA or SSE) (Shanks, 2013) and (2) provide evidence for genetic differentiation between benthic-stage Dungeness crab from the CCE and northern ecosystems (Jackson and O’Malley, 2017; O’Malley et al., 2017). Furthermore, we identified two putatively adaptive loci, for which strong genetic differentiation was observed between expected-season and late-season recruits. Taken together, these findings suggest that late-season megalopae recruits may have originated from a northern ecosystem.

Although differentiation between expected-season and late-season megalopae recruits was significant based on variation at presumably neutral loci, the $F_{ST}$ estimate (0.0011) was recognizably low, even for a marine species. Gene flow is expected to be high for marine species given that few dispersal barriers exist and effective population sizes are typically large (Palumbi, 1994; Ward et al., 1994; Waples, 1998). However, significant yet low $F_{ST}$ estimates have been found to be biologically meaningful in several genomic studies of marine species with lengthy PLDs. For example, a study of southern rock lobster (Jasus edwardsii) larval recruits concluded the significant inter-annual genetic differentiation between years and within sites resulted from ocean currents influencing the annual sources of larvae ($F_{ST}$ range: 0.0062–0.0240; Villacorta-Rath et al., 2017). Moreover, a study on the giant California sea cucumber (Parastichopus californicus) demonstrated that species with a longer PLD that span the divide between northern and southern ecosystems along the North American west coast may exhibit lower, but biologically relevant $F_{ST}$ estimates ($F_{ST} = 0.002$) that have implications for management of the species (Xuereb et al., 2018). Taken in the context of these past studies, we conclude that more years of data are needed to confirm if our finding of weak genetic differentiation between expected- and late-season recruits is biologically meaningful.

Heterozygote deficiency (positive $F_{IS}$) is often observed in marine invertebrates, and the reasoning for this deficiency is often population or study specific (reviewed in Raymond et al., 1997; Addison and Hart, 2005). Positive $F_{IS}$ can be attributed to several processes: selection, inbreeding, laboratory or data artifacts, null alleles, or the Wahlund effect (i.e., unrecognized spatial or temporal structure within samples) (Addison and Hart, 2005). The Wahlund effect describes a situation in which there is a deficit of heterozygotes due to the pooling of several breeding groups into one sample (Wahlund, 1928). Since reproducing Dungeness crab are typically confined to a small spatial area due to limited benthic movement (Hildenbrand et al., 2011), the departure from HWP may result from a mixture of larvae from a small number of parental cohorts. The Wahlund effect has been observed often in marine species (Planes and Lenfant, 2002).

The GBS approach used in this study allowed for the identification of over a thousand loci including two putatively adaptive loci, where past genetic studies on Dungeness crab used 10 neutral microsatellite loci (Jackson and O’Malley, 2017; Jackson et al., 2017; O’Malley et al., 2017). Using both neutral and adaptive loci together to examine genomic variation within a species is advantageous because it provides information on gene flow and genetic drift, as well as natural selection. For
example, in a study on the adult stage of the eastern rock lobster (*Sagmariasus verreauxi*) (8–12 month PLD) within the Tasman Sea, population structure was not observed based on variation at 645 neutral loci, but based on variation at 15 outlier loci (high $F_{ST}$) informative genetic differentiation was observed between geographic regions (Woodings et al., 2018). None of the 15 outlier loci mapped to genes, but the researchers presumed the loci may represent adaptive regions. We observed strong differentiation between the expected-season and late-season recruits based on variation at the two putatively adaptive loci. This finding may be attributed to differences in selective pressures between the CCE and the more northern SSE and GOA, such as temperature and ocean chemistry. Such selective pressures would result in adaptive differences between groups of individuals originating from different regions. It is important to emphasize that these putatively adaptive loci are outliers with high $F_{ST}$ values, and although informative to questions of differentiation, may not truly reflect regions of the genome under selection (Shafer et al., 2015). However, one of the putatively adaptive loci identified between megalopae recruits exhibited similarity to a hypothetical mRNA sequence in the NCBI database for the Asian mud crab (*S. paramamosain*), providing additional support that the identified loci may be adaptive. It is unsurprising that only one putatively adaptive locus matched because, currently, there are a limited number of mapped and annotated crustacean genomes (Rotllant et al., 2018).

The benthic stage CCE Dungeness crab have exhibited inter-annual variations in population genetic structure ranging from weak IBD (2012) to panmixia (2014) based on variation at neutral microsatellite loci (Jackson et al., 2017). Presumably, the 2014 panmictic population would have contributed to the 2014 larval cohort; Dungeness crab reach sexual maturity at ~100 mm width carapace (Rasmuson, 2013). With a panmictic reproductive population in the CCE, we would expect the larval recruits of 2014 to be genetically homogenous if they all originated from within the CCE. However, we observed weak differentiation between the expected-season and the late-season megalopae recruits based on variation at presumably neutral loci and stronger differentiation based on putatively adaptive loci. This finding supports the idea that the late-season megalopae recruits did not originate from within the CCE.

Alternatively, the late-season megalopae recruits could be offspring of a group of CCE adults reproducing later in the season than documented or the PLD may be longer than 3–4 months for some Dungeness crab within the CCE. Therefore, the megalopae recruits would not need to be from different ecosystems (CCE vs. SSE or GOA) to be genetically differentiated. This alternative scenario could still represent a different reproductive source of megalopae leading to weak, yet significant, intra-annual differentiation.

Further evidence is needed to confirm the hypothesis that the source of the late-season megalopae recruits is from an ecosystem north of the CCE. Such evidence would include a comparison of the late-season megalopae recruits to reproducing adult populations in the CCE, SSE, and GOA. Additionally, a reproducible pattern of intra-annual differentiation across multiple years and/or a stronger pairwise $F_{ST}$ estimate between expected- and late-season recruits based on variation at the presumably neutral loci could indicate that recruits originate from a different ecosystem. Past population genetic studies on Dungeness crab have only observed weak genetic differentiation among adult Dungeness crab in the CCE at neutral loci (Jackson et al., 2017; O’Malley et al., 2017), so a strong signature of genetic differentiation would suggest that one of the megalopae recruitment cohorts did not originate from within the CCE. GBS approaches utilizing both neutral and adaptive loci can provide greater power to detect fine-scale population structure than previous microsatellite studies (Vendrami et al., 2017).

**Ocean Conditions and Megalopae Recruitment in 2014**

In an effort to understand how ocean conditions allow for the exchange of Dungeness crab larvae within and between ecosystems, we examined the 2014 megalopae recruitment abundance within the context of previous years.
can improve our understanding of this important fishery. Techniques to study the early life stage of the Dungeness crab fishery, this case study demonstrates that utilizing genomic data from the coupled human-natural system of the Dungeness crab megalopae recruits collected from a site in Coos Bay every year monitored prior to 2018 (1998-2001 and 2007-2017) of daily recruitment monitoring (Shanks and Roegner, 2007; Shanks et al., 2010; Shanks, 2013, 2018). In 2014, the abundance of late-season recruits was below average but above the median, ranking as the ninth lowest year of megalopae recruitment in a 15-year dataset (1998-2001 and 2007-2017) of daily recruitment monitoring (Shanks and Roegner, 2007; Shanks et al., 2010; Shanks, 2013; unpublished data Shanks, 2018).

Based on models by Shanks and Roegner (2007); Shanks et al. (2010), and Shanks (2013), the lower abundance of recruiting megalopae in 2014 should have resulted in a lower commercial catch during the 2018 Dungeness crab commercial fishery season. However, the Dungeness crab commercial catch within the CCE during the 2018 fishery season was above average and the sixth largest seasonal catch in reported history (California Department of Fish and Wildlife, 2018; Oregon Department of Fish and Wildlife, 2018; Washington Department of Fish and Wildlife, 2018). It is uncertain as to why more Dungeness crab were caught in the 2018 commercial fishery than expected by megalopae abundance; however, Shanks (2018) has suggested that the atypical ocean warming event from 2014 to 2016 (i.e., “the Blob,” Peterson et al., 2017) may have played a role in this discrepancy. This highlights the fact that researchers do not fully understand the relationship between megalopae recruitment, adult population size, and, ultimately, the fishery harvest. Therefore, a better understanding of how megalopae recruitment corresponds to adult population size is needed to understand the implications of early life stage variations for the fishery, especially in the face of changing ocean conditions. In this study, we demonstrated that there are genetic differences between expected- and late-season larval recruits in the CCE. This information may be applicable to future management efforts assessing the adaptive capacity of Dungeness crab in the CCE to changing ocean conditions.

CONCLUSION

In this paper, we present and interpret genomic data from Dungeness crab megalopae recruits collected from a site within the CCE where megalopae recruitment patterns have been documented and studied for over two decades. Although there are still many uncertainties surrounding the coupled human-natural system of the Dungeness crab fishery, this case study demonstrates that utilizing genomic techniques to study the early life stage of the Dungeness crab can improve our understanding of this important fishery species through enhanced knowledge of the fine-scale, intra-annual genetic differentiation between expected-season and late-season megalopae recruits. Furthermore, the laboratory and bioinformatic methods provide a framework for future genomic studies of Dungeness crab megalopae. This case study of the 2014 Coos Bay, Oregon Dungeness crab megalopae recruitment cohort highlights the importance of understanding early life stage variation, as it provides a 4-year foresight into the socioeconomically important Dungeness crab fishery.

DATA AVAILABILITY STATEMENT

This manuscript contains previously unpublished data. The GENEPOP file containing all filtered loci (1,915) is available through Dryad, doi: https://doi.org/10.5061/dryad.7wm37pvpb.

ETHICS STATEMENT

No vertebrate animal or cephalopod studies are presented in this manuscript. All applicable international, national, and institutional guidelines for the care and use of animals were followed.

AUTHOR CONTRIBUTIONS

EL and KO designed the study, analyzed the data, and wrote the manuscript.

FUNDING

This study was funded by NSF-NRT award #1545188 (Risk and uncertainty quantification and communication in marine science and policy), the Oregon Sea Grant Robert E. Malouf Marine Studies Scholarship (grant #NA14OAR4170064 and project #E/INT-157), and the Hatfield Marine Science Center Mamie Markham Research Award. The statements, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of these funders.

ACKNOWLEDGMENTS

We thank Alan Shanks and the members of the Shanks lab at the Oregon Institute of Marine Biology, University of Oregon for providing the megalopae samples and light trap abundance data for this study. We also thank Oregon State University’s Center for Genome Research and Biocomputing for preparing and sequencing our megalopae samples. Thanks to our team of faculty and students in the Oregon State University National Science Foundation Research Traineeship (NRT) program, Risk and Uncertainty Quantification in Marine Science for supporting this project. And special thanks to the members of the State Fisheries Genomics Lab at Hatfield Marine Science Center for providing valuable guidance throughout this project.
Paris, J. R., Stevens, J. R., and Catchen, J. M. (2017). Lost in parameter space: a road map for stacks. *Methods Ecol. Evol.* 8, 1360–1373. doi: 10.1111/2041-210x.12775

Pemberton, L. W., Cogan, N. O., and Forster, J. W. (2013). STAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Mol. Ecol. Res.* 13, 946–952. doi: 10.1111/1755-0998.12129

Peterson, W. T., Fisher, J. L., Strub, P. T., Du, X., Risien, C., Peterson, J., et al. (2017). The pelagic ecosystem in the Northern California Current Off Oregon during the 2014–2016 warm anomalies within the context of the past 20 years. *J. Geophys. Res. Oceans* 122, 7267–7290. doi: 10.1002/2017jc012952

Planes, S., and Lenfant, P. (2002). Temporal change in the genetic structure between and within cohorts of a marine fish, *Diplodus sargus*, induced by a large variance in individual reproductive success. *Mol. Ecol.* 11, 1515–1524. doi: 10.1046/j.1365-294x.2002.01521.x

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795

R Core Team, (2018). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rasmussen, L. K. (2013). “The biology, ecology and fishery of the Dungeness crab, *Cancer magister*,” in *Advances in Marine Biology*, Vol. 65, ed. M. Lesser (Cambridge, MA: Academic Press), 95–148. doi: 10.1016/b978-0-12-410498-3.00003-3

Raymond, M., Väntö, R. L., Thomas, F., Roussel, F., de Meûts, T., and Renaud, F. (1997). Heterozygote deficiency in the mussel *Mytilus edulis* species complex revisited. *Mar. Ecol. Prog. Ser.* 156, 225–237. doi: 10.3354/meps156225

Ritzman, J., Ritzman, J., Brud, M., McGrew, S., Dreyer, S., Klinger, T., et al. (2017). RAD sequencing resolves fine-scale population structure in a benthic invertebrate: implications for understanding phenotypic plasticity. *R. Soc. Open Sci.* 4:160548. doi: 10.1098/rsos.160548

Villacorta-Rath, C., Souza, C. A., Murphy, N. P., Green, B. S., Gardner, C., and Strugnell, J. M. (2017). Temporal genetic patterns of diversity and structure evidence chaotic genetic patchiness in a spiny lobster. *Mol. Ecol.* 27, 54–65. doi: 10.1111/mec.14427

Wahlund, S. (1928). The combination of populations and the appearance of correlation examined from the standpoint of the study of heredity. *Hereditas* 11, 65–106.

Waples, R. S. (1998). Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *J. Hered.* 89, 438–450. doi: 10.1093/jhered/89.5.438

Ward, R. D., Woodmark, M., and Skibinski, D. O. (1994). A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. *J. Fish. Biol.* 44, 213–232. doi: 10.1111/j.1095-8649.1994.tb01200.x

Washington Department of Fish and Wildlife, (2018). Available at: https://wdfw.wa.gov/ (Accessed December 1, 2018).

Weir, B. S., and Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Am. J. Hum. Genet.* 83, 747–760. doi: 10.1093/amjhg/83.3.747

Whitlock, M. C., and Lotterhos, K. E. (2015). Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of FST. *Am. Nat.* 186, S24–S36.

Wigginton, J. E., Cutler, D. J., and Abecasis, G. R. (2005). A note on exact tests of Hardy-Weinberg equilibrium. *Am. J. Hum. Genet.* 76, 887–893. doi: 10.1086/429864

Wild, P. W., and Tasto, R. N. (1983). *Life History, Environment, and Mariculture Studies of the Dungeness Crab, Cancer magister, with Emphasis on the Central California Fishery Resource*. Sacramento, CA: Department of Fish and Game California.

Woodings, L. N., Murphy, N. P., Doyle, S. R., Hall, N. E., Robinson, A. J., Liggins, G. W., et al. (2018). Outlier SNPs detect weak regional structure against a background of genetic homogeneity in the Eastern Rock Lobster, *Sagmariasus verreauxi*. *Mar. Biol.* 12, 165–185.

Xuereb, A., Benestan, L., Normandeau, E., Daigle, R. M., Curtis, J. M., Bernatchez, L., et al. (2018). Asymmetric oceanographic processes mediate connectivity and population genetic structure, as revealed by RADseq, in a highly dispersive marine invertebrate (*Parastichopus californicus*). *Mol. Ecol.* 27, 2347–2364. doi: 10.1111/mec.14589

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Lee and O’Malley. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.