Title: **An Initial Comparative Genomic Autopsy of Wasting Disease in Sea Stars**

Mobile device title: Genomic autopsy of wasting disease

Dannise V. Ruiz-Ramos[a,*,1], Lauren M. Schiebelhut[a,*], Katharina J. Hoff[b,c], John P. Wares[d], Michael N Dawson[a]

[a]Department of Life and Environmental Sciences, 5200 North Lake Road, University of California, Merced, CA 95616, USA
[b]University of Greifswald, Institute for Computer Science and Mathematics, Walther-Rathenau-Str. 47, 17487 Greifswald, Germany
[c]University of Greifswald, Center for Functional Genomics of Microbes, Felix-Hausdorff-Str. 8, 17489 Greifswald, Germany
[d]Department of Genetics and the Odum School of Ecology, University of Georgia, Athens, GA, USA

*Equal first authors

[1]To whom correspondence should be addressed. Dannise V. Ruiz-Ramos, School of Natural Sciences, University of California, 5200 N. Lake Rd., Merced, CA 95343, Email: dannise.ruiz@gmail.com.

Keywords: disease, gene expression, genome, mass mortality, sea star wasting disease (SSWD), asteroid idiopathic wasting syndrome (AIWS)

Abstract

Beginning in 2013, sea stars throughout the Eastern North Pacific were decimated by wasting disease, also known as 'asteroid idiopathic wasting syndrome' (AIWS) due to its elusive etiology. The geographic extent and taxonomic scale of AIWS meant events leading up to the outbreak were heterogeneous, multifaceted, and oftentimes unobserved; progression from morbidity to death was rapid, leaving few tell-tale symptoms. Here we take a forensic genomic approach to discover candidate genes that may help explain sea star wasting syndrome. We report the first genome and annotation for *P. ochraceus*, along with differential gene expression (DGE) analyses in four size classes, three tissue types, and in symptomatic and asymptomatic individuals. We integrate nucleotide polymorphisms associated with survivors of the wasting disease outbreak, DGE associated with temperature treatments in *P. ochraceus*, and DGE associated with wasting in another asteroid *Pycnopodia helianthoides*. In *P. ochraceus*, we find DGE across all tissues, among size classes, and between asymptomatic and symptomatic individuals; the strongest wasting-associated DGE signal is in pyloric caecum. We also find previously identified outlier loci co-occur with differentially expressed genes. In cross-species comparisons of symptomatic and asymptomatic individuals, consistent responses distinguish genes associated with invertebrate innate immunity and chemical defense, consistent with context-dependent stress responses, defensive apoptosis, and tissue degradation. Our analyses thus highlight genomic constituents that may link suspected environmental drivers (elevated temperature) with intrinsic differences among individuals (age/size, alleles associated with susceptibility) that elicit organismal responses (e.g. coelomocyte proliferation) and manifest as sea star wasting mass mortality.

Introduction

Wildlife mass mortality events are increasing in frequency, driven in part by human perturbations, emerging viral diseases, intensifying toxic algal blooms, increasing temperatures, or a combination of multiple stressors (Fey et al., 2015). In the marine environment, disease emergence may be accelerated by environmental change, as warming and acidification can suppress the host immune response (Alker, Smith, & Kim, 2001; R. W. Chapman et al., 2011; Harvell, Altizer, Cattadori, Harrington, & Weil, 2009) and favor the growth and spread of pathogens (R. W. Chapman et al., 2011; Harvell et al., 2009). Novel terrestrial microbes may be introduced to coastal waters where they encounter new hosts or change their pathogenicity (Alker et al., 2001; Burge et al., 2014; Harvell et al., 2009). Marine infectious diseases, whether novel or endemic, have caused mass mortality events (MMEs) in diverse taxa, including plants (Burge, Kim, Lyles, & Harvell, 2013; Short, Muehlstein, & Porter, 1987), invertebrates (Burge, Griffin, & Friedman, 2006; Gardner et al., 1995; Lessios, 2016) and vertebrates (Dubey et al., 2003) of commercial and ecological importance.

Some events contributed to the collapse of fisheries (Castro, Cobb, Gomez-Chiarri, & Tlusty, 2012) and ecosystems (Lessios, 2016); the MME of the keystone species *Pisaster ochraceus* associated with sea star wasting disease (SSWD) caused top-down trophic cascades in subtidal reefs (Schultz, Cloutier, & Côté, 2016) and altered zonation of the rocky intertidal community (Gravem & Morgan, 2017). Most of these MMEs occurred without forewarning; causes, mechanisms, and effects had to be inferred *post hoc* in the absence of information about pre-MME conditions. While our abilities to reconstruct likely factors from genomic signatures (De Wit, Rogers-Bennett, Kudela, & Palumbi, 2014) and infer longer-term decline of vulnerable

70 populations with genomic tools (Bay et al., 2018) are improving, we remain largely ignorant of
71 genomic attributes that may elevate susceptibility or resilience to MMEs.
72 The 2013 outbreak of SSWD in the northeastern Pacific Ocean was remarkable for several
73 reasons, including that both ecological and population genetic sampling were underway
74 preceding the outbreak (Harley, Pankey, Wares, Grosberg, & Wonham, 2006; Miner et al., 2018;
75 Schiebelhut, Puritz, & Dawson, 2018). This enabled documentation of the outbreak's greatest
76 geographic extent (Eisenlord et al., 2016), highest mortality rates recorded for a non-commercial
77 marine species—e.g. 67% (Eisenlord et al., 2016), 90% (Menge et al., 2016), 99% (Miner et al.,
78 2018)—and diversity of sea star species affected (Eisenlord et al., 2016; Montecino-Latorre et al.,
79 2016). Over 20 species of subtidal and intertidal asteroids were impacted (Hewson et al., 2014),
80 of which some species, like *Pisaster ochraceus* and *Pycnopodia helianthoides*, are important
81 predators in their communities (Duggins, 1983; Gravem & Morgan, 2017; Paine, 1974). Thus,
82 mass mortality of sea stars precipitated cascading effects through intertidal and subtidal
83 communities (Burt et al., 2018; Menge et al., 2016; Schultz et al., 2016).
84 Given the large geographic extent (Miner et al., 2018) and taxonomic scale (Eisenlord et
85 al., 2016; Montecino-Latorre et al., 2016) of SSWD, events leading up to the outbreak were
86 heterogeneous, multifaceted, and often unrecognized at the time (Miner et al., 2018). Moreover,
87 the disease status of specimens in the field progresses rapidly from apparently healthy through
88 morbid to dead, leaving few opportunities for diagnoses prior to, or in the presence of, visually
89 identifiable symptoms (Miner et al., 2018), so clarifying evidence has been elusive. Proposed
90 causes are diverse, including infection, injury, and environmental shifts (Harvell et al., 2019;
91 Hewson et al., 2018). Differences in susceptibility also have been noted, possibly related to
92 age/size (Eisenlord et al., 2016; Menge et al., 2016), temperature (Eisenlord et al., 2016; Harvell

et al., 2019; Kohl, McClure, & Miner, 2016), habitat (Menge et al., 2016), region (Miner et al., 2018), and genotypic variation (Schiebelhut et al., 2018; Wares & Schiebelhut, 2016), or more speculatively salinity (Bates, Hilton, & Harley, 2009). Nonetheless, causes and effects remain obscure, and SSWD has recently been rebranded as asteroid idiopathic wasting syndrome (AIWS) due to its elusive etiology (Hewson et al., 2018).

Despite coincidence of the outbreak of wasting (i.e. SSWD/AIWS) with formation of a northeastern Pacific 'warm blob' (Bond, Cronin, Freeland, & Mantua, 2015), the role of temperature in disease onset and intensification is ambiguous (Hewson et al., 2018; Miner et al., 2018), suggesting that wasting might be caused by a combination of environmental (Hewson et al., 2018; Miner et al., 2018) and biological factors (Menge et al., 2016). Exploration of differential gene expression in *P. ochraceus* under thermal stress suggested heritable variation in transcriptional response to elevated temperature (Chandler & Wares, 2017). Viral inoculation experiments with a sea star associated densovirus (SSaDV) found symptoms of wasting after exposure and showed wasting to be transmissible in the sunflower star, *Pycnopodia helianthoides* (Hewson et al., 2014). However, SSaDV was also found in asymptomatic individuals (Hewson et al., 2018, 2014), and subsequent inoculation experiments failed to induce wasting symptoms in species other than *P. helianthoides* (Hewson et al., 2018). Analyses of symptomatic and asymptomatic *P. helianthoides* found differential gene expression in genes associated with immune response, nervous system processes (Fuess et al., 2015), tissue disintegration (Gudenkauf & Hewson, 2015), tissue remodeling (Fuess et al., 2015), programmed cell death, and metabolic function (Gudenkauf & Hewson, 2015).

In association with the high rates of mortality associated with wasting, analyses of single-nucleotide polymorphism (SNP) data from *P. ochraceus* populations along the north-central coast

of California identified consistent genotypic differences between adult *P. ochraceus* before versus after the 2013 wasting outbreak (Schiebelhut et al., 2018). Changes in allele frequency after the mass mortality were largely consistent across locations, and between life stages, suggesting that wasting acted as a strong selective event (Schiebelhut et al., 2018).

To address some of the remaining knowledge gaps between observations of mortality and genomic responses, we here present (i) a high-quality reference genome and (ii) new gene expression data from multiple life stages, tissue types, and disease statuses of the ochre sea star *Pisaster ochraceus*. We then (iii) compare these data to other published gene expression studies, including a second sea star *P. helianthoides*, and interrogate the genome for clues on the etiology of wasting. Our goal is to discover genomic signals that could tie-together the somewhat eclectic ecological and experimental results available in the literature. Particularly, we follow four lines of enquiry to attempt to reconstruct plausible scenarios leading to wasting. We first narrow down how gene expression differs between tissue types, ages (i.e. sizes), and disease status. Second, we explore the extent to which expression responses to disease overlap with expression responses to one proposed cause: temperature. We then ask whether expression differences implicate specific immune pathways or are concordant with responses to wasting across species. Finally, we ask which, if any, of the implicated gene regions also showed allele frequency shifts in wasting-survivors, possibly indicating a link between mutations in coding/noncoding sequence and gene expression. Identification of genomic elements that respond in common across species or conditions should provide clues to the causes and conditions that increase the risk of wasting (SSWD/AIWS).

Materials and Methods

Genome sample collection, sequencing, and assembly

An asymptomatic *Pisaster ochraceus* of 49 mm radius (arm-tip to disc-center; sample ID: M0D055189C) was collected from the intertidal zone at Hopkins Marine Station on 12 November 2016, and taken alive to Dovetail Genomics (Santa Cruz, CA, USA) where tube feet were amputated and frozen immediately in liquid nitrogen. DNA extraction, library construction (paired-end, Chicago, HiC), Illumina sequencing, and *de novo* assembly followed Dovetail Genomics standard approaches (J. A. Chapman et al., 2011; Lieberman-Aiden et al., 2009; Putnam et al., 2016). Sequences were deposited in the Genome database of the National Center for Biotechnology Information (NCBI; BioProject PRJNA532896, SUB5448653). See SI Appendix for details. To assess completeness of the assembly, we conducted a BUSCO v. 2.0.1 search against the set of metazoan orthologous genes (Simão, Waterhouse, Ioannidis, Kriventseva, & Zdobnov, 2015).

Transcriptome sequencing

Tissue biopsies—of tube feet, pyloric caecum, and dermis—were taken from 17 individuals: 15 collected at McClures Beach and two from Lifeboat House in Point Reyes National Seashore (SI Appendix, Table S1) on the 19th and 21st August 2017, respectively, and preserved in RNA*later*® in the field, refrigerated overnight, and stored at -20°C until extraction. RNA was extracted with the Qiagen RNeasy Mini Kit. A Retsch mixer mill (MM 301) was used for 30 seconds for tissue disruption and homogenization. Fifty RNA libraries were generated (SI Appendix, Table S1) with Illumina TruSeq RNA Library Prep Kit v2 and sequenced 2 x 150 paired-end on one lane of an Illumina HiSeq4000 with 2% Phi X at the Vincent J. Coates Genomics Sequencing Laboratory at the University of California, Berkeley. Raw sequences were

deposited in the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI; add SRA accession number).

Genome annotation

Repetitive regions of the genome assembly were identified and softmasked using RepeatMasker (http://www.repeatmasker.org) to align the genome to the RepBase repository (RepeatMasker Edition 20170127; https://www.girinst.org/server/RepBase) (Bao, Kojima, & Kohany, 2015) using *Drosophila* and *Strongylocentrotus purpuratus* as the species options. As an alternative, the genome was also soft-masked using NCBI WindowMasker (Morgulis, Gertz, Schäffer, & Agarwala, 2005). The RepeatMasker masking with *S. purpuratus* was used to aid annotation of protein coding genes.

Gene models were generated with BRAKER v 2.1.0 (Hoff, Lange, Lomsadze, Borodovsky, & Stanke, 2015; Hoff, Lomsadze, Borodovsky, & Stanke, 2019), a pipeline that combines the gene prediction tools GeneMark-ET v. 4.3.2 (Lomsadze, Burns, & Borodovsky, 2014) and AUGUSTUS v. 3.3 (Stanke, Diekhans, Baertsch, & Haussler, 2008) for predicting genes in novel genomes with alignments of RNA-Seq reads as extrinsic evidence. The 50 RNA libraries were supplemented with paired-end RNA libraries from Chandler and Wares (Chandler & Wares, 2017) prepared from individuals with distinct genotypic backgrounds (SAMN06141149, SAMN06141151) to capture a full range of transcript data. In short, RNA-Seq libraries were aligned to the soft masked genome using GSNAP v. 2017-11-15 (Wu, Reeder, Lawrence, Becker, & Brauer, 2016), and BRAKER was run on the soft masked genome with information from spliced alignments (for code, see Data Accessibility section). A track data hub

for annotation visualization with the UCSC Genome Browser (Kent et al., 2002) was generated using MakeHub (Hoff, 2019). The software faCount (UCSC) was used to count base composition in each scaffold. Percentages of gaps—regions that were not sequenced or gaps included during scaffolding—were calculated from these data (faCount) by dividing the number of Ns in each chromosome-sized scaffold by the total number of bases in the scaffold. GC content in contiguous windows of 500 nt was calculated using the GC function in the SeqinR package (Charif & Lobry, 2007; Coghlan, 2011) and a logic function in Excel to count the regions with >60% of bases being C+G.

We used NCBI Blastp v. 2.2.31+ (Altschul, Gish, Miller, Myers, & Lipman, 1990; Camacho et al., 2009) to annotate the predicted gene models by searching invertebrate protein sequences from the SwissProt protein database version 2018_02 (Uniprot Consortium, 2014; "UniProt: The universal protein knowledgebase," 2016). The most representative annotation from UNIPROT ID was taken as the ID with lowest e-value, unless there was >1 lowest match in which case we favored the UNIPROT ID from the closest relative. The table of the corresponding ID numbers was downloaded to add gene ontology (GO; Ashburner et al., 2000; The Gene Ontology Consortium, 2016)(GO v. 2018_02) terms to the annotation (Dataset S1) (for code, see Data Accessibility section). GO terms were then grouped by categories using GO_slim and, given our interest in disease, the immune system gene classes, with CateGOrizer (Hu, Bao, & Reecy, 2008)—a GO term classification counter—and redundant parent terms were summarized with REViGO (Supek, Bosnjak, Skunca, & Smuc, 2011).

Mitochondrial genome

Raw reads generated for the whole genome assembly (section *Chicago library preparation and sequencing* in SI Appendix) were aligned to the annotated mitochondrial

genome of the confamilial *Asterias amurensis* (NC_006665; (Matsubara et al., 2005)) to generate a whole mitochondrial assembly. A subsample of $3x10^7$ reads were assembled using the mapping function with "high sensitivity" parameters in Geneious 10.1.2, adjusted to accept reads with 25% mismatch. Annotations were assessed manually from the alignment of this assembly with the *A. amurensis* mitogenome, and via submission to MITOS (Bernt et al., 2013). Additional annotation, including the location of the mitochondrial control region and adjustments to the computationally estimated ORF of *ND2*, came from pairwise alignment with the mitochondrial genome of *Acanthaster planci* (Yasuda et al., 2006).

Differential gene expression

We added a second lane of Illumina 2 x 150 paired-end HiSeq4000 sequencing with 2% Phi X for the previously mentioned 50 RNA libraries (see *Transcriptome sequencing* section) to increase depth and coverage of transcripts for differential gene expression (DGE) analyses. Libraries represented individuals that were asymptomatic or symptomatic for SSWD (e.g. Fig. 1) and of different size classes—a proxy for age: small (radius 13.0–17.1 mm), medium (43–67 mm), large (89–123 mm), and extra-large (165–212 mm)—for each of three tissue types (tube feet, dermis, and pyloric caecum) (SI Appendix, Table S1). Raw sequences were again deposited in the NCBI SRA (accession no. SRPXXXXXX). Trim Galore v. 0.4.4_dev was used to trim adapters, low quality bases (<Q20), and reads less than 20bp.

Genestack v. 0.72.1 (https://genestack.com/) was used for analyses of trimmed sequences. Trimmed paired-end reads were subsampled to 5 million reads (smallest library size) and aligned to the *Pisaster ochraceus* nuclear genome using TopHat2 v. 2.0.13 (Kim et al., 2013) allowing 2 mismatches per read and yielding mean 2,599,568 mapped reads per library (range: 2,313,438–2,969,060). (See SI Appendix Supplementary Methods and Results for alternative mapping

approach). For the mitochondria, trimmed R1 reads were aligned to the mitochondrial genome using Bowtie2 v. 2.2.3 (Langmead, Trapnell, Pop, & Salzberg, 2009) allowing 3 mismatches per read; reads were then subsampled to the smallest mapped library (0.213 million reads) using seqtk v. 1.0 (https://github.com/lh3/seqtk.git), yielding 0.213 million mapped reads per library. Raw coverage in genes[1] was quantified using SAMtools v. 0.1.19 (Li et al., 2009). Differential gene expression was tested using EdgeR (McCarthy, Chen, & Smyth, 2012) for sequences aligned to the nuclear genome and mitochondrial genome separately. A DGE analysis was first conducted to evaluate differential expression between tissue types. Subsequent analyses were conducted separately for each tissue type to assess differential expression between size classes and between symptomatic and asymptomatic individuals. Each size class (i.e. small, medium, large, or x-large) was compared to the mean of all other size classes. An additional analysis was conducted comparing just the small individuals directly to the x-large individuals to reveal potential DGE that might be masked by variation in intermediate size classes. Genes were filtered at a minimum threshold of mean 1 count per million (CPM). Significance was assessed at a false discovery rate (FDR) < 0.01 (unless otherwise noted); power to detect DGE will be influenced by the number of reads in libraries. Given the small sample sizes of individuals included in the asymptomatic (n = 2) compared to symptomatic (n = 2) analyses (SI Appendix, Table S1), a minimum criterion of CPM > 1 in a minimum of two libraries per asymptomatic and/or symptomatic categories was used. Gene enrichment analyses were performed in DAVID for the

[1] Gene models were defined, for complete genes, as beginning at the start codon and ending at the stop codon. For incomplete genes, gene models were recognized as spanning from the 5' to 3' sequence boundaries within the gene or, if partially incomplete, then between the gene boundary (start/stop codon) and matching sequence boundary.

annotated differentially expressed genes (Huang, Sherman, & Lempicki, 2008, 2009) using the list of *Pisaster ochraceus* annotated genes as background.

Reanalysis of DGE experiments

Raw sequences from Chandler and Wares' (Chandler & Wares, 2017) *P. ochraceus* temperature experiment—in which tube feet were sampled from 10 individuals acclimated to ambient ocean temperature and resampled after 8 days at +3°C—were obtained from the Sequence Read Archive (SRA) at NCBI (SRP095092). Forward sequences were reprocessed and analyzed following the method described above in the section *Differential gene expression*. Each filtered RNA library was subsampled to 22.5 million reads (the smallest library) and then mapped to the nuclear genome, yielding mean 16,325,675 mapped reads (range: 15,830,301–16,754,055). For the mitochondria, all forward reads were mapped to the mitochondrial genome and then filtered to the smallest mapped library at 1.15 million reads. A DGE analysis contrasted between individuals at ambient ocean temperature (~12°C with diurnal variation, lab acclimated for 2 days, 't0') and +3°C (tissue collected after 8 days of treatment, 't3'). Genes were filtered at a minimum threshold of 1 CPM. Significant difference between the t0 and t3 groups was assessed at a false discovery rate (FDR) < 0.1. Because Chandler and Wares (2017) biopsied the same 10 individuals before and after heat treatment, we also conducted a two-tailed paired *t* test ($p < 0.05$ following Benjamini-Hochberg adjustment for multiple tests; https://www.sdmproject.com/utilities/?show=FDR) contrasting time-points within individuals to identify other potential genes that might be differentially expressed under temperature stress but not detected by comparing just the two groups. We filtered these results further by conducting a χ^2 goodness-of-fit test ($p < 0.05$) to focus on genes that had consistent directional change in expression between t0 and t3 across individuals (i.e. when ≥9 of 10 individuals agreed).

For mitochondrial analyses, raw sequences from Fuess et al. (Fuess et al., 2015) *P. helianthoides* symptomatic versus asymptomatic experiment—in which coelomocytes were sampled from 3 individuals (with wasting symptoms) after being treated with virus-sized homogenate taken from wasting *P. helianthoides* and 3 control stars were treated with heat-killed homogenate (and showed no signs of wasting)—were obtained from the NCBI SRA (SRP051104). Forward sequences were reprocessed and aligned to the *P. ochraceus* mitochondrial genome and analyzed following the method described above in the section *Differential gene expression*. Each filtered and mapped RNA library was subsampled to 60,900 mapped reads (the fewest number of mapped reads of any individual). A DGE analysis contrasted symptomatic and asymptomatic individuals using a minimum threshold of 1 CPM and FDR < 0.1. Raw reads from Gudenkauf & Hewson (2015) did not yield enough read depth to assess mitochondrial DGE in *P. helianthoides*.

Mapping and comparison of genomic and transcriptomic variation

We used the annotated *Pisaster ochraceus* genome and DGE analyses conducted in this study to map and compare findings from other asteroid studies, including: (1) RAD-Seq loci putatively under selection by wasting (Schiebelhut et al., 2018), (2) genes differentially expressed between *Pycnopodia helianthoides* symptomatic and asymptomatic for wasting (Fuess et al., 2015; Gudenkauf & Hewson, 2015), and (3) genes differentially expressed in *P. ochraceus* undergoing temperature and/or related stress in aquarium experiments (Chandler & Wares, 2017). Particularly, we aimed to infer chromosomes or regions of chromosomes that may be of heightened interest for understanding the causes and effects of SSWD, as follows. (1) To evaluate potential function of the top RAD-Seq loci discriminating pre- and post- AIWS (Schiebelhut et al., 2018), the position of each locus in relation to the gene models was inspected visually, and

295 the transcripts that showed differential expression between individuals asymptomatic and
296 symptomatic for AIWS were noted as falling inside or outside of the gene model region. To
297 assess mitochondrial allele frequency changes between pre- (n = 153) and post-wasting adults (n
298 = 126) and recruits (n = 158) we aligned RAD-Seq reads from Schiebelhut et al. (Schiebelhut et
299 al., 2018) (NCBI SRA SRP136569) to the *P. ochraceus* mito-genome; 8 individuals were
300 dropped for low read count and an additional 3 dropped for ambiguous base calls. (2) To evaluate
301 gene position and function, RNA-Seq reads from coelomocytes and dermis punches of
302 *Pycnopodia helianthoides* symptomatic and asymptomatic for AIWS (Fuess et al., 2015;
303 Gudenkauf & Hewson, 2015) were mapped to the *Pisaster ochraceus* genome with BWA-MEM
304 v. 0.7.15 (Li & Durbin, 2009), sorted with SAMtools v. 1.4.1 (Li et al., 2009), merged and
305 converted to a bed file with BEDTools v. 2.23.0 (Quinlan & Hall, 2010). Finally, (3) transcripts
306 differentially expressed in *P. ochraceus* at ocean ambient and +3°C temperature treatments from
307 the re-analysis of Chandler and Wares (Chandler & Wares, 2017) (see section *Temperature*
308 *treatment comparisons*) were compared to the other DGE analyses in this study. The mapped
309 positions and/or functions of loci from all studies were compared to identify shared patterns
310 across datasets and so infer loci that may be important in the etiology of AIWS in *P. ochraceus*
311 and possibly also other sea stars. The R package ggbio (Yin, Cook, & Lawrence, 2012) was used
312 to visualize these results in concentric circular maps to highlight the positions of loci implicated
313 to be common components in the etiology of wasting disease.
314 In the absence of any general theoretical expectations for the number of loci and their
315 effect sizes influencing complex phenotypes such as disease susceptibility (in sea stars), we
316 modelled the probability of a specific locus occurring in more than one analysis using a simple
317 probability series. For this purpose, we estimated the probability of a locus being implicated in a

single analysis as the 'number of DGE loci / gene count' for that individual analysis. The probability of the same locus being implicated by multiple analyses is then the product of the individual probabilities; the probabilities for all possible combinations of 1–5 analyses are provided in Dataset S9. Furthermore, the probability that the locus is differentially expressed in the same direction is 0.5^n, where n is number of analyses considered. We also conducted a Fisher's Exact Test in R, corrected for multiple tests using the Benjamini-Hochberg procedure (https://www.sdmproject.com/utilities/?show=FDR), to calculate the probability of overlap between genes found to be differentially expressed between *P. ochraceus* symptomatic and asymptomatic for sea star wasting disease and other studies — i.e. *P. ochraceus* size, temperature treatment, ddRAD outliers, and *Pycnopodia helianthoides* disease status.

Non-random distribution of RAD-Seq and DGE loci—using super-scaffold length and the length adjusted by the number of Ns in the scaffolds (Fig. 1) as null models for predicted distributions—was tested using a χ^2 test. The outcomes did not differ qualitatively so only results relative to the adjusted lengths are reported.

Results

Genome

Shotgun sequencing generated 551 million paired-end reads (totaling 78 Gbp) which were assembled *de novo* into a draft assembly, to which Chicago library sequences were added to produce scaffolds up to 13 Mb (N50 = 8.8 kb, total genome size = 366.6 Mb). Dovetail HiC library sequences were aligned and scaffolded, and gaps filled with shotgun sequences to create 22 super scaffolds of approximately chromosome size (N50= 21 Mb) and a total genome size of 401.95 Mb (Fig. 1). The total percent of genome in gaps was 13.23% (range 10.0–20.5% per

chromosome-size scaffold). GC content of 500 nt windows ranged from 1.6% to 92.6% (mean 39.1%, median 38.9%; SI Appendix, Fig. S1) with 860 regions having GC content >60%.

A BUSCO comparison before masking indicated the *P. ochraceus* genome had 88.1% genes complete, 9.6% genes fragmented, and 2.3% genes missing (SI Appendix, Fig. S2). 24% of the genome was masked as areas of high repeat content first with WindowMasker (Morgulis et al., 2005). RepeatMasker (Chen, 2004) with *Drosophila* and *S. purpuratus* as species option, identified 1.35% of the masked genome as simple repeats and 0.38% as transposable elements, mostly Non-LTR SINE2 repeats.

21,777 genes were predicted for the *Pisaster ochraceus* genome using the BRAKER pipeline. 98% of the gene models were found in the 22 super scaffolds; the number of gene models was correlated with super-scaffold size ($R^2 = 0.804$). Only 7,861 (36%) had significant hits (e-value < 0.00001) to the non-redundant invertebrate protein sequences from the SwissProt protein database (UniProt Consortium, 2016) (Dataset S2) and 7,694 (35%) gene models matched one or more (up to 145) GO terms, totaling 853,637 GO terms (Dataset S2). The GO terms mapped to 108 (of 110) immune class ancestors terms (CateGOrizer); the main fractions were metabolism (39%), stress response (13%), and protein metabolism (11%); the classes 'stress response', 'response to external stimulus', response to 'abiotic stimulus', 'apoptosis', and 'response to biotic stimuli' also were in the top ten most numerous classes, with the 'mitochondrion' class 12th, the 'immunology, immune response' class is 13th, and 'defense response to bacteria' is 24th (Dataset S3a, SI Appendix, Fig. S3). Twenty-one genes matched Toll-like receptors and Leucine-rich repeat-containing proteins (Dataset S3b), fifteen of these clustered within the Toll signaling pathway (Dataset S3c). Genes associated with the regulation of innate immune response (n=34), stress response (n=37) and humoral response (n=42) were

364 identified (Dataset S3c). Some gene models were also annotated as retrovirus-related Pol
365 polyproteins from LRT retrotransposons (Dataset S3d), which were most abundant in
366 chromosome 12 (34 gene models).
367
368 *Mitochondrial genome*
369 The full mitochondrial genome (NCBI accession MH713001) was recovered from the
370 unassembled sequences. The mitochondrial genome is 16,376 bp in length, with all recognized
371 protein, ribosomal, and tRNA genes (SI Appendix, Fig. S4).
372
373 *Differential gene expression*
374 *Tissue comparisons* — Most differential gene expression (FDR < 0.01) was identified when
375 comparing different tissue types (SI Appendix, Fig. S5). Approximately 5,000 genes were
376 differentially expressed among tissues: 5,481 in pyloric caecum (of which 2,160 were annotated),
377 5,046 in tube feet (2,092 annotated), and 4,584 in the dermis (1,954 annotated) (SI Appendix,
378 Fig. S6). Due to this large difference in expression, other targeted differential expression analyses
379 were done separately for each tissue type.
380 *Size comparisons* — Gene expression analyses comparing size classes (i.e. small, medium, large,
381 and x-large versus all others, and smallest versus x-largest) revealed 261 nuclear genes were
382 differentially expressed (FDR < 0.01) among size classes in the pyloric caecum, 295 in the
383 dermis, and 414 in the tube feet (Fig. 2; Dataset S4; SI Appendix, Fig. S7). The majority of these
384 genes were differentially expressed in small individuals versus other size classes—88% of genes
385 in pyloric caecum, 62% in dermis, and 90% in tube feet. Restricting the analysis to the 3 smallest
386 (radius 13.0–17.1 mm) and 3 x-largest (165–212 mm) individuals resulted in 177 genes

387 differentially expressed in the pyloric caecum (72 of which were annotated), 75 genes in the
388 dermis (20 annotated), and 227 in the tube feet (90 annotated) (Fig. 2, Dataset S5). When
389 comparing smallest to x-largest individuals we find upregulation in genes associated with
390 hemolymph coagulation (in the dermis), lipid biosynthesis (pyloric caecum; fold enrichment =
391 6.4, $p < 0.001$, FDR = $7x10^{-4}$; Dataset S12), lipid metabolism (pyloric caecum; fold enrichment =
392 5.1, $p < 0.001$, FDR = $2x10^{-5}$; Dataset S12) and metabolism (dermis and tube feet) (SI Appendix,
393 Fig. S8A). Most processes were upregulated in the dermis of small stars, including those related
394 to development, ovarian (fold enrichment = 9.7, $p = 0.03$, FDR = 0.34; Dataset S12) and
395 metabolic processes (fold enrichment = 1.4, $p = 0.03$, FDR = 0.38; Dataset S12) (SI Appendix,
396 Fig. S8A). In the tube feet, cell differentiation (fold enrichment = 4.7, $p = 0.05$, FDR = 0.54;
397 Dataset S12), vesicle-mediated transport, actin capping, and thermotaxis were upregulated in
398 small stars (SI Appendix, Fig. S8A).
399 In analyses of the mitochondria, DGE analyses revealed higher expression of *tRNA-Gln*
400 (in pyloric caecum) and *tRNAs-Asp, Asn*, and *Met* (in tube feet) of small individuals relative to
401 other size classes, while *16S L-rRNA, ATP6, COX2* were downregulated in small individuals
402 relative to other size classes. In the dermis of medium individuals, *16S L-rRNA* was upregulated
403 relative to other groups, while *16S L-rRNA* in the tube feet of x-large individuals was upregulated
404 relative to small individuals. Medium individuals also had increased expression of *ND4L* in their
405 dermis relative to other size classes.
406 *Symptomatic vs. asymptomatic comparisons* — Nuclear genes differentially expressed between
407 symptomatic and asymptomatic samples were found in the pyloric caecum (581 genes, 257
408 annotated), dermis (259 genes, 94 annotated) and tube feet (49 genes, 23 annotated) (Fig. 3;
409 Dataset S6).

In the pyloric caecum, 346 genes (148 annotated) were upregulated in the symptomatic individuals, while 235 genes (109 annotated) were downregulated (Fig. 3; Dataset S6a). 35 of the genes differentially expressed between symptomatic-asymptomatic individuals were associated with immune or stress response (Dataset S6a, Dataset S3c)—such as genes involved in hemocyte proliferation (*TL5A*, *TL5B*), agglutination and phagocytosis (*SRCR1*)—and an additional four involved in programmed cell death (fold enrichment = 1.6, $p = 0.09$; Dataset S12)—e.g. apoptosis (*CASP1*) and programmed cell death (e.g. *CED1*) (Fig. 4, Dataset S3c). Toll-like genes putatively involved in the Toll signaling pathway (fold enrichment = 9.5, $p = 0.04$, FDR = 0.46; Dataset S12) were upregulated in the symptomatic individuals (Dataset S3c). Enriched GO terms also included lipid metabolism (fold enrichment = 1.6, $p = 7\text{x}10^{-4}$; Dataset S12), GDP-fucose biosynthesis, catabolism and metabolism (steroid: fold enrichment = 4.6, $p = 0.003$, FDR = 0.05; monosaccharide: fold enrichment = 4.5, $p = 0.02$, FDR = 0.31; Dataset S12). Additionally, ten genes of the Cytochrome P450 family (*CP10* = g2103; *CP18A* =g17638, g6090, g6091; *CP2L1* = g7145, g9883; *CP4D2* = g12454, g12447; *CP9E2* = g4034; *C524A* = g18059) were differentially expressed in the pyloric caecum of the symptomatic samples (Dataset S6a).

In the dermis, 166 genes were upregulated in the symptomatic samples (58 annotated) and 93 genes were downregulated (36 annotated) (SI Appendix Fig. S8B, Dataset S6b). 14 differentially expressed genes were associated with defense response (fold enrichment = 3.8, $p = 0.04$, FDR = 0.48; Dataset S12), wound healing (response to wounding: fold enrichment = 6.6, $p = 0.02$, FDR = 0.28; Dataset S12), apoptosis (*CASP1*) and clotting processes (Dataset S3c; Dataset S6b). As in the pyloric caecum, genes involved in the regulation of immune response (immune system processes: fold enrichment = 2.7, $p = 0.04$, FDR = 0.51; Dataset S12) and the Toll signaling pathway were upregulated in symptomatic individuals (Dataset S3c). Genes

associated with antimicrobial humoral response and hemocyte proliferation (*TL5A, TL5B, LEC6*) were also downregulated in the dermis of the symptomatic individuals (Dataset S3c). Genes associated with wound healing were upregulated in the symptomatic samples, however, *Cadherin 96Ca* which is involved in positive regulation of wound healing was downregulated. Enriched GO terms for negative regulation of innate immune response (fold enrichment = 6.27, $p = 0.01$, FDR = 0.11; Dataset S3c; Fig. 4) and recognition of apoptotic cells were upregulated in the symptomatic samples (SI Appendix, Fig. S8B).

In the tube feet, 37 genes were upregulated in the symptomatic samples (19 annotated) and 12 genes were downregulated (4 annotated) (Fig. 3; Dataset S6c). Genes expressed in tube feet were mainly associated with muscular function (fold enrichment = 15.7, $p = 0.01$, FDR = 0.17) and cell adhesion (Fig. 4; SI Appendix, Fig. S8B), but two genes associated with immune response *TL5A* and *NOXC* were upregulated in symptomatic samples (Dataset S6c).

Individuals symptomatic for AIWS (n = 2), in DGE analyses of the mitochondria, had higher expression (FDR < 0.01) of *12S s-rRNA* in all three tissue types independently—dermis (fold change of 3.3, FDR < 10^{-20}), pyloric caecum (fold change of 20.6, FDR = 10^{-5}), and tube feet (fold change of 2.6, FDR = $4x10^{-4}$)—relative to asymptomatic stars (n = 2). In the dermis, *tRNA-Asn* and *16S L-rRNA* mitochondrial genes were also upregulated in symptomatic, relative to asymptomatic stars (FDR = 0.009 and 0.063, respectively).

Reanalysis of DGE experiments — Genes differentially expressed between the tube feet of sea stars sampled at ambient ocean temperature and after heat treatment (+3°C) included 39 identified by the DGE analysis in EdgeR (FDR < 0.1) and 137 identified by the paired *t* test followed by filtering for similar directional change in ≥9 of 10 individuals ($\chi^2 \geq 6.40$, df = 1, N = 10, $p < 0.05$); 59 of the 137 were significant after Benjamini-Hochberg adjustment $p_{0.05}$. 35 loci

were recovered by both analyses. The EdgeR analysis (FDR < 0.1) comparing the t0 group to the t3 group also recovered four transcripts that did not overlap with the paired *t* test, three of which were unannotated (g4748, g10034 and g20429) and one annotated (g12264)—the heat shock protein *Hsp90*. In total, 73 loci were identified as being significantly differentially expressed by one or both of these approaches.

In DGE analyses of the mitochondria, individuals after heat treatment (n = 10) had lower expression (FDR < 0.1) of ATP Synthase subunit 6 (*ATP6*) and NADH dehydrogenase subunits 2 and 5 (*ND2*, *ND5*) and increased expression (FDR < 0.1) of *16S L-rRNA* and *tRNA-Asp* relative to t0 (SI Appendix, Fig. S9).

In mitochondrial analyses of coelomocytes in *P. helianthoides*, no differential expression was found between symptomatic (n = 3) and asymptomatic individuals (n = 3).

Mapping and comparison of genomic and transcriptomic variation

Comparisons of multiple studies—RAD-Seq, DGE analyses of *P. ochraceus* by disease status, size, and temperature treatment, and *P. helianthoides* by disease status—revealed shared and distinct genomic responses (Fig. 5; Dataset S7). Of the top 100 RAD-Seq loci identified in Schiebelhut et al. (2018) discriminating pre- and post-AIWS mortality populations, 99 mapped to the 22 chromosomes in the *P. ochraceus* genome (Fig. 5); 58 aligned to predicted gene models (Dataset S8), 25 of which were annotated. Five loci mapped to genes that were differentially expressed between *P. ochraceus* symptomatic and asymptomatic for SSWD (FDR < 0.01) (SI Appendix, Fig. S10; Fig. 5): hypothetical protein *vilA* (g3886), mitochondrial putative tricarboxylate transport protein *K11H3.3* (g17130), GTP-binding protein *drn-1* (g21294), and unannotated g868 were all upregulated in symptomatic individuals, while unannotated g7633 was downregulated. *vilA* also was upregulated in the coelomocytes of SSWD-symptomatic *P.*

helianthoides (Fuess et al., 2015) as well as overlapping with one of the three RAD-Seq loci identified in the more conservative BayeScan analysis as discriminating pre- and post-SSWD *P. ochraceus* (Schiebelhut et al., 2018). Four additional discriminant RAD-Seq loci (Schiebelhut et al., 2018) mapped to genes upregulated in SSWD-symptomatic *P. helianthoides* (Fuess et al., 2015): E3 ubiquitin-protein ligase *Nedd-4* (g14570), muscle M-line assembly protein *unc-89* (g19459), Rho GTPase-activating protein 92B *RhoGAP92B* (g20941), and one unannotated gene (g20271) (Dataset S7b). One discriminant RAD-Seq locus mapped to homeobox protein *Lhx3* (g19194), a gene upregulated in the tube feet of small individuals (Dataset S7b).

Thirty-three transcripts upregulated in SSWD-symptomatic individuals in Gudenkauf & Hewson (2015) mapped to *P. ochraceus* gene models (SI Appendix, Table S3; Dataset S7a), of which six were also differentially upregulated in SSWD-symptomatic *P. ochraceus*: lateral signaling target protein 2 *lst2* (g18551), frizzled *Dvir\fz* (g3737), and four unannotated gene models (g3818, g6204, g14951, and g18769) (Dataset S7c,d).

Immune and neurological DGE responses of *Pycnopodia helianthoides* to treatment with a viral sized homogenate (Fuess et al., 2015) yielded 1387 differentially expressed transcripts that mapped to the *P. ochraceus* gene models (Fig. 5; Dataset S7a). There was considerable overlap between DGE of symptomatic *P. ochraceus* and symptomatic *P. helianthoides*—112 of 129 genes (86.8%) that overlapped between the studies shared similar differential expression (SI Appendix, Table S3; Dataset S7c,d)

DGE responses to heat treatment of *P. ochraceus* (Chandler & Wares, 2017) yielded 73 differentially expressed transcripts that mapped to the *P. ochraceus* genome (Fig. 5; Dataset S7a). Of these, 6 transcripts overlapped with genes differentially expressed between symptomatic and asymptomatic *P. ochraceus* (Dataset S7c,d). Three of these genes were upregulated in both heat-

502 treated and AIWS-symptomatic sea stars—Vitellogenin-6 (*vit-6*; g13155), and two unannotated
503 genes (g10034, g12499). Three other genes were downregulated—*Hyalin* (g2770), Spectrin beta
504 chain (*beta-Spec*; g9020), and one unannotated (g13988).
505 Comparison of DGE in symptomatic/asymptomatic *P. ochraceus* and size classes reveal
506 12.6% (114/902) of genes from the size analysis overlap with genes differentially expressed in
507 SSWD-symptomatic individuals, and of those 59% (67/114) share up/down-regulation in smaller
508 (relative to larger individuals) with symptomatic (relative to asymptomatic) individuals (SI
509 Appendix, Table S3; Dataset S7a). A subset of 9 of these genes overlap with two or more other
510 studies (Fig. 5A; Dataset S7d), four of which show similar regulation in smaller individuals and
511 symptomatic *P. ochraceus* and *P. helianthoides*—upregulation in Dopamine transporter (*DAT*,
512 g12628) and Myosin heavy chain (*MYS*, g21191), downregulation in Dehydrogenase/reductase
513 SDR family member 4 (*dhrs-4*, g2110) and WW domain containing oxidoreductase (*Wwox*,
514 g21298)—and one that is downregulated in smaller individuals, symptomatic *P. ochraceus*, and
515 temperature-treated *P. ochraceus* (Spectrin beta chain (*beta-Spec*; g9020). Four genes show the
516 opposite pattern—Ets at 98B (*Ets98B*, g7334) and Fibrinogen-like protein A (*FIBA*, g9029) are
517 downregulated in smaller individuals, but upregulated in symptomatic *P. ochraceus* and *P.*
518 *helianthoides*, while Putative fatty acid elongation protein 3(*elo-3*, g5517) and Organic cation
519 transporter 1(*oct-1*, g9945) are upregulated in smaller individuals, but downregulated in
520 symptomatic *P. ochraceus* and *P. helianthoides* (Dataset S7d). An additional 8 transcripts
521 differentially expressed between size classes were also differentially expressed in heat-treated *P.*
522 *ochraceus* (Dataset S7b), with shared upregulation in smaller (versus larger) and heat-treated
523 individuals of Actin (*ACT1*, g11029) and (g4110, no annotation), and shared downregulation in
524 alpha Spectrin (*alpha-Spec*, g9023). Four of the eight transcripts were downregulated in smaller

individuals, but upregulated in heat-treated *P. ochraceus*—N-acetylglucosamine-6-phosphate deacetylase (*CG17065*, g16112) and 3 unannotated genes (g7624, g14020, g19130)—while one unannotated gene (g14019) was upregulated in smaller individuals, but downregulated in heat-treated *P. ochraceus*.

In addition to the nuclear analysis, we compared mitochondrial responses and found *16S L-rRNA* had increased expression in both temperature treated *P. ochraceus* (SI Appendix, Fig. S9) and AIWS-symptomatic *P. ochraceus* (Fig. 5B). We also identified a synonymous substitution in *ND5* in 3.2% (n = 4 of 125) of surviving adults and 2.5% (n = 4 of 154) of juveniles that was undetected in the pre-wasting population of *P. ochraceus* and this gene was downregulated in temperature treated stars.

The probability of a specific locus being implicated in multiple analyses ranges from 0.0039 in 2 analyses, through ~0.0007–0.0001 in 3 analyses, to ~0.00003–0.00001 in 4 analyses. The probability that these implicated loci also have the same direction of differential gene expression is even smaller at 10^{-4} in 2 analyses, through 10^{-5} in 3 analyses, to 10^{-6} –10^{-7} in 4 analyses. A Fisher's Exact Test, corrected for multiple tests using the Benjamini-Hochberg procedure, supports overlap in genetic mechanisms associated with wasting disease in *P. ochraceus* and those associated with *P. ochraceus* size ($p < 10^{-15}$) and *P. helianthoides* disease ($p = 0.002$, $p < 10^{-15}$); Fisher's Exact Test is marginally non-significant for overlap between *P. ochraceus* disease-associated differentially expressed genes ($p = 0.06$) and temperature-associated differentially expressed genes or disease-associated ddRAD loci ($p = 0.06$).

The outlier loci described above were distributed across all or almost all chromosomes. The top 100 RAD-Seq loci of Schiebelhut et al. (2018) were randomly distributed among chromosomes ($\chi^2 = 26.00$, df = 21, N = 99, $p = 0.205$), as were the RNA-seq loci differentially

548 expressed between symptomatic and asymptomatic individuals (χ^2= 32.39, df = 21, N = 802, p =
549 0.053). However, RNA-seq loci differentially expressed after heat treatment were non-randomly
550 distributed, with a notable cluster of 8 loci on the fourth shortest Chromosome 19 (χ^2 = 35.57, df
551 = 21, N = 73, p = 0.024), as were loci that were differentially expressed by size, which were
552 disproportionately numerous on chromosomes 16 and 18–20 (χ^2= 84.30, df = 21, N = 897, $p < 10^{-8}$).
553
554
555 **Discussion**
556 For the first time since the 2013 outbreak of sea star wasting disease (SSWD)—one of the largest
557 marine mass mortality events on record—we are able to integrate data from field observations,
558 field collections, and laboratory experiments on multiple species within the common framework
559 of a high quality reference genome. This physical genetic map links a suspected environmental
560 driver (elevated temperature) and intrinsic differences among individuals (age/size, genotype) and
561 genes underlying organismal responses (e.g. immune response [*TL5A*, *TL5B*], phagocytosis
562 [*SRCR1*], cell death and wound healing [*CASP1*], apoptosis [*WWOX*], muscle contraction [*MYS*],
563 heat shock protein [*HSP71*]) with presentation of SSWD. These results help define what was
564 becoming known as asteroid idiopathic wasting syndrome (AIWS; (Hewson et al., 2018)) due to
565 its previously elusive etiology. We do not claim to circumscribe the entire set of factors involved
566 in SSWD, but to have identified important candidate genes and to have described an approach—
567 an autopsy guided by genomic analyses—for better understanding causes and mechanisms of
568 MMEs. Through a set of logical steps, our preliminary exploration suggests several candidate
569 genetic mechanisms for further investigation of the relationship between the sea star wasting

570 outbreak and large scale environmental change are merited (see also Fey et al., (2015); Harvell et
571 al., (2019)).

572 The genome and annotation of *Pisaster ochraceus* is the first for Order Forcipulatida, a
573 diverse and ecologically notable clade of sea stars (Mah & Blake, 2012). The assembly is ~90%
574 complete with 22 super-scaffolds, matching the haploid number of chromosomes estimated by
575 karyotyping (Saotome & Komatsu, 2002). At ~402 Mb, the *P. ochraceus* genome tends toward
576 the middle of the size-spectrum for most animals and deuterostomes (Canapa, Barucca, Biscotti,
577 Forconi, & Olmo, 2015; Elliott & Gregory, 2015). The modest size of the *P. ochraceus* genome
578 suggests relatively few transposable elements (Elliott & Gregory, 2015). Its GC content (mean
579 ~39.1%) is marginally higher than that of *S. purpuratus* (~36.9%; (Sodergren et al., 2006)), and
580 the 0.1% of 500 nt windows with >60% CG content may be interesting to investigate in the
581 context of methylation status and potential CpG islands proximate to genes implicated in wasting.
582 The estimated number of genes in *P. ochraceus* is marginally higher than predicted for a genome
583 of this size (Elliott & Gregory, 2015). The potential conservation suggested here—of
584 chromosome number, genome size, and gene number within the clade—merits further
585 investigation, as does genome architecture, in the context of wasting as an asteroid zoonosis.

586 Our new analyses of differential gene expression in *P. ochraceus* reveal differences
587 among tissues, among sizes, and between symptomatic versus asymptomatic individuals. The
588 greatest differences reflect the distinct expression needed during development to generate
589 different tissues (Ralston & Shaw, 2008) and then to deliver their complementary functions.
590 These functions have recently become of particular interest given the putative involvement of
591 densoviruses in wasting (Hewson et al., 2014; but see Hewson et al., 2018), that the microbiome
592 of sea stars appears to be anatomically partitioned (Jackson, Pepe-Ranney, Debenport, Buckley,

593 & Hewson, 2018), and that tissues may display different prevalence of viruses within and
594 between species (Hewson et al., 2018). The pyloric caecum shows the greatest number of
595 transcripts being differentially expressed between symptomatic and asymptomatic *P. ochraceus*
596 (Fig. 3; Fig. 4; SI Appendix, Fig. S8B); the cause is unclear but may help explain the prior
597 association between SSaDV and wasting: elevated transcription could indicate hyperplasia, and
598 densoviruses are favored by rapidly dividing cells (Tijssen, Pénzes, Yu, Pham, & Bergoin, 2016).
599 The dermis, where lesions occur most prominently, also has a high degree of differential
600 expression between symptomatic and asymptomatic *P. ochraceus*. Less affected, though not
601 unaffected, are the tube feet, consistent with field observations that even very sick sea stars still
602 attach to rocks and separated arms walk. Likewise, differential gene expression within tissues
603 between sizes (Fig. 2), common in many organisms, is also consistent with field observations that
604 differences in susceptibility are possibly related to age/size (Eisenlord et al., 2016; Menge et al.,
605 2016).

606 Differential gene expression within tissues between symptomatic versus asymptomatic
607 individuals occurs in both mitochondrial and nuclear genes. One mitochondrial locus (*12S s-*
608 *rRNA*) was upregulated in all three tissue types in symptomatic *P. ochraceus* (relative to
609 asymptomatic individuals). The 12S ribosomal region intriguingly is a focus of recent studies
610 examining the role of mitochondria in (human) systemic disease and apoptosis (Raimundo et al.,
611 2012; Shadel, 2008). This may be meaningful for understanding wasting and its linkage with
612 apoptosis (Fig. 4; Dataset S6; Dataset S7d; Dataset S8; (Fuess et al., 2015; Gudenkauf &
613 Hewson, 2015)). One mitochondrial gene (*16S L-rRNA*) was upregulated in symptomatic and
614 heat-treated *P. ochraceus*. Heat-treated stars showed upregulation of another mitochondrial gene
615 (*tRNA-Asp*) and downregulation of three mitochondrial genes (*ATP6*, *ND2*, *ND5*), relative to stars

kept at ambient ocean temperature (SI Appendix, Fig. S9). The *ND* subunits appear relevant because decreased activity in the mitochondrial oxidative phosphorylation (OXPHOS) electron transport chain leads to increased production of reactive oxygen species (ROS) (Kamogashira, Fujimoto, & Yamasoba, 2015), which can lead to increased apoptotic cell death (Zapico & Ubelaker, 2013) and contribute to mitochondrial permeabilization, triggering caspase-induced apoptosis (Oberst, Bender, & Green, 2008; Tait & Green, 2010). Commensurately, DGE analysis of the nuclear genome evinced caspase activation in the symptomatic stars: *Caspase-1*, which is involved in the activation cascade of caspases responsible for apoptosis execution, was elevated in the symptomatic samples, both in the pyloric caecum and dermis. Intriguingly, in humans, *Caspase-1* also regulates the response to DNA viruses upon inflammasome activation (Wang et al., 2017). Other nuclear genes that were differentially expressed in all tissues in the symptomatic–asymptomatic comparison included those associated with immune defense, cell adhesion and wound healing. Humoral (*TLs-5)* and cellular (*SRCR)* components of the asteroid innate immune system (Hibino et al., 2006; Ramírez-Gómez & García-Arrarás, 2010) were represented in the differentially expressed genes in the symptomatic versus asymptomatic comparisons. In symptomatic individuals genes associated with activation of the immune response (*KPEL*, g15346; *CED1*, g19731) were upregulated; for coelomocyte proliferation, different transcripts of *SRCR1* were upregulated (g2511, g12941) or downregulated (g234, g6871), and agglutination agents downregulated (*TLs-5*, g14108; *LEC6*, g2945) in different tissues of symptomatic samples. SRCR genes are expressed specifically in the coelomocytes (Pancer, 2000), cells that are abundant and diverse in echinoderms, with roles in chemotaxis, clotting, phagocytosis and encapsulation (Ramírez-Gómez & García-Arrarás, 2010). Due to their multiple functions, it is unsurprising that different SRCR1 genes have different expression

639 patterns. The downregulation of lectins (*TLs-5* and *LEC6*) in the pyloric caecum and dermis of
640 symptomatic stars, suggests wasting is not of bacterial origin: TLs-5 [*TL5A* (g14108) and *TL5B*
641 (g18833)] are lectins involved in innate immunity and antimicrobial activity functioning in non-
642 self recognition (Gokudan et al., 1999) and strong agglutinating response to bacteria (Gokudan et
643 al., 1999). Additionally, we found differential gene expression between symptomatic and
644 asymptomatic *P. ochraceus* at loci putatively involved in selection (Fig. S10) associated with
645 SSWD (Schiebelhut et al., 2018), including upregulation of GTP-binding protein (*drn-1*, g21294)
646 in individuals symptomatic for SSWD, suggesting there could be a link between genotypic
647 variation and gene expression associated with SSWD.
648 We acknowledge our survey of currently available genomic data may miss loci of weaker
649 effect, loci in the unassembled 13% of the genome, and loci not involved in multiple pathways,
650 that influence SSWD. Additionally, *cis*-regulatory evolution can play an important role in gene
651 expression that may be important in adaptation (Lasky et al., 2014), although our current analyses
652 do not address this explicitly. Nonetheless, our approach highlights candidate loci that provide a
653 common thread tying together the ecological observations reported to date. Statistically, some of
654 these candidates may be false positives; however, the probability that any particular locus is
655 implicated as a candidate locus purely by chance is small ($p = \sim10^{-4}–10^{-7}$; Fisher's Exact Test, $p =$
656 $0.06 – p < 10^{-15}$) and so we consider the highlighted loci to be of considerable interest.
657 One of the most intriguing aspects of the 2013 outbreak of sea star wasting disease has
658 been its wide zoonotic impact and yet apparently few consistent responses among species in
659 subsequent studies (e.g., see Hewson et al. (2018) and Miner et al. (2018)). Our synthesis of prior
660 datasets (Chandler & Wares, 2017; Fuess et al., 2015; Gudenkauf & Hewson, 2015; Schiebelhut
661 et al., 2018) enabled by the common reference of the *P. ochraceus* genome however begins to

suggest some potential commonalities (Fig. 5), although we do not consider the role of complex interactions among stressors, which may manifest in unintuitive ways that do not necessarily result in shared differential expression of the same genes. However, we do find evidence linking DGE responses to temperature treatments with DGE in symptomatic (versus asymptomatic) individuals. Additionally, we find a synonymous substitution in mitochondrial *ND5* in adult and juvenile survivors that was undetected in pre-wasting populations. All sea stars exposed to increased temperature showed a decrease in *ND5* expression (Fig. 5, SI Appendix, Fig. S9) relative to stars at ambient ocean temperature. And while synonymous (Goymer, 2007) mutations in *ND5* can lead to oxidative phosphorylation disease (Blok et al., 2007), there are many polymorphisms in mtDNA, including in the control region, that could affect transcription. These results are particularly interesting as recent long-term coast-wide analyses have suggested a link between temperature and the wasting disease outbreak of 2013 (Eisenlord et al., 2016; Harvell et al., 2019; Kohl et al., 2016). Nonetheless, regional stressors may vary (Hewson et al., 2018), and whether wasting is a direct response to temperature stress or is part of a general stress response that increases disease risk, is still unknown (Miner et al., 2018). That the differentially expressed genes for temperature treated stars are distributed across all 22 chromosomes and the 6 genes that overlap with symptomatic versus asymptomatic DGE are distributed across 6 different chromosomes (Dataset S7d), does suggest that the sea star response is a general response (i.e. not attributable to a single gene or few genes), fitting its innate-only immune system. The potential conservation of genomic attributes and organismal responses may therefore indicate underlying similarities that could help explain the susceptibility of over 20 species of subtidal and intertidal asteroids to wasting (Hewson et al., 2014). To resolve this question, controlled laboratory

684 experiments will be needed, along with expansion of the comparative genomic approach to
685 include many more asteroid species with different reported susceptibilities to wasting.
686 The temporal rapidity and geographic heterogeneity of the 2013 sea star wasting disease
687 outbreak made it difficult to narrow down the range of factors, causes, and mechanisms involved
688 at the time. Similar to other multi-taxon pandemics (e.g. “white nose” syndrome in bats, chytrid
689 fungi for amphibians, morbillivirus in dolphins) SSWD was elusive in the early stages of study
690 (Hewson et al., 2018; Miner et al., 2018), which is a concern as mass mortalities are increasing in
691 frequency with global change in some taxa (Fey et al., 2015; Tracy et al., 2019) and require more
692 rapid assessment. Given the central role of asteroid predators in community ecology (Gravem &
693 Morgan, 2017; Menge, 1983; Paine, 1974), as the focal organisms in the 2013 wasting pandemic,
694 and wasting’s intermittent recurrence (albeit with different intensity and taxonomic breadth) we
695 desperately need to better understand risk factors. While our abilities to reconstruct potential
696 agents from genomic signatures (De Wit et al., 2014) and infer longer-term decline of vulnerable
697 populations with genomic tools (Bay et al., 2018) are improving, we often remain ignorant of
698 genomic attributes that may elevate susceptibility or resilience to MMEs. Our approach—
699 retrospectively comparing “omic” changes across taxa, using the *Pisaster* genome to arrange
700 expression and other data associated with disease status and other phenotypes, i.e. genomic
701 autopsy—could form a model for approaching other emerging diseases. This approach can help
702 us understand genomic attributes that may shape—sometimes very differently, as in *Pisaster* and
703 *Pycnopodia*—organismal, population, species, and ecosystem responses to, and in the aftermath
704 of, marine MMEs (Burge et al., 2016; Lafferty & Hofmann, 2016).
705
706 **Acknowledgements**

Funding was provided by a University of California President's Research Catalyst Award (CA-16-376437) and the National Science Foundation Biological Oceanography program (OCE-1737381, OCE-1737091). The California Department of Fish and Wildlife and National Park Service provided permits and access to field sites. Computation time was provided by the MERCED cluster at UC Merced, funded by National Science Foundation award ACI-1429783. We thank Ian Hewson for discussion of disease etiology and viruses and members of the University of California Conservation Genomics Consortium (especially the Wayne lab at UCLA and Shapiro lab at UCSC) for discussion of genomics. The Raimondi lab at UCSC kindly assisted with collections.

References

Alker, A. P., Smith, G. W., & Kim, K. (2001). Characterization of *Aspergillus sydowii* (Thom et Church), a fungal pathogen of Caribbean sea fan corals. *Hydrobiologia*, *460*(1–3), 105–111.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*, 403–410.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., … Eppig, J. T. (2000). Gene Ontology: Tool for the unification of biology. *Nature Genetics*, *25*, 25–29.

Bao, W., Kojima, K. K., & Kohany, O. (2015). Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile Dna*, *6*(1), 11.

Bates, A. E., Hilton, B. J., & Harley, C. D. (2009). Effects of temperature, season and locality on wasting disease in the keystone predatory sea star *Pisaster ochraceus*. *Diseases of Aquatic Organisms*, *86*(3), 245–251.

Bay, R. A., Harrigan, R. J., Le Underwood, V., Gibbs, H. L., Smith, T. B., & Ruegg, K. (2018). Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science*, *359*(6371), 83–86.

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., … Stadler, P. F. (2013). MITOS: improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, *69*(2), 313–319.

Blok, M. J., Spruijt, L., de Coo, I. F. M., Schoonderwoerd, K., Hendrickx, A., & Smeets, H. J. (2007). Mutations in the ND5 subunit of complex I of the mitochondrial DNA are a frequent cause of oxidative phosphorylation disease. *Journal of Medical Genetics*, *44*(4), e74–e74.

Bond, N. A., Cronin, M. F., Freeland, H., & Mantua, N. (2015). Causes and impacts of the 2014 warm anomaly in the NE Pacific. *Geophysical Research Letters*, *42*(9), 3414–3420.

743 Burge, C. A., Eakin, C. M., Friedman, C. S., Froelich, B., Hershberger, P. K., Hofmann, E. E., … Willis,
744 B. L. (2014). *Climate change influences on marine infectious diseases: Implications for*
745 *management and society*.
746 Burge, C. A., Friedman, C. S., Getchell, R., House, M., Lafferty, K. D., Mydlarz, L. D., … Kiryu, I.
747 (2016). Complementary approaches to diagnosing marine diseases: A union of the modern and the
748 classic. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1689),
749 20150207.
750 Burge, C. A., Griffin, F. J., & Friedman, C. S. (2006). Mortality and herpesvirus infections of the Pacific
751 oyster *Crassostrea gigas* in Tomales Bay, California, USA. *Diseases of Aquatic Organisms*,
752 *72*(1), 31–43.
753 Burge, C. A., Kim, C. J. S., Lyles, J. M., & Harvell, D. (2013). Special Issue Oceans and Humans Health:
754 The ecology of marine opportunists. *Microbial Ecology*, *65*(4), 869–879.
755 Burt, J. M., Tinker, M. T., Okamoto, D. K., Demes, K. W., Holmes, K., & Salomon, A. K. (2018). Sudden
756 collapse of a mesopredator reveals its complementary role in mediating rocky reef regime shifts.
757 *Proceedings of the Royal Society B: Biological Sciences*, *285*(1883), 20180553.
758 Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009).
759 BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*(1), 421.
760 Canapa, A., Barucca, M., Biscotti, M. A., Forconi, M., & Olmo, E. (2015). Transposons, genome size, and
761 evolutionary insights in animals. *Cytogenetic and Genome Research*, *147*(4), 217–239.
762 Castro, K. M., Cobb, J. S., Gomez-Chiarri, M., & Tlusty, M. (2012). Epizootic shell disease in American
763 lobsters *Homarus americanus* in southern New England: Past, present and future. *Diseases of*
764 *Aquatic Organisms*, *100*(2), 149–158.
765 Chandler, V. K., & Wares, J. P. (2017). RNA expression and disease tolerance are associated with a
766 “keystone mutation” in the ochre sea star *Pisaster ochraceus*. *PeerJ*, *5*, e3696.

767 Chapman, J. A., Ho, I., Sunkara, S., Luo, S., Schroth, G. P., & Rokhsar, D. S. (2011). Meraculous: De
768 novo genome assembly with short paired-end reads. *PloS One*, *6*(8), e23501.
769 Chapman, R. W., Mancia, A., Beal, M., Veloso, A., Rathburn, C., Blair, A., … Sokolova, I. M. (2011).
770 The transcriptomic responses of the eastern oyster, *Crassostrea virginica*, to environmental
771 conditions. *Molecular Ecology*, *20*(7), 1431–1449.
772 Charif, D., & Lobry, J. R. (2007). SeqinR 1.0-2: A contributed package to the R project for statistical
773 computing devoted to biological sequences retrieval and analysis. In *Structural approaches to*
774 *sequence evolution* (pp. 207–232). Springer.
775 Chen, N. (2004). Using RepeatMasker to identify repetitive elements in genomic sequences. *Current*
776 *Protocols in Bioinformatics*, *5*(1), 4.10. 1-4.10. 14.
777 Coghlan, A. (2011). *Little book of R for bioinformatics*.
778 De Wit, P., Rogers-Bennett, L., Kudela, R. M., & Palumbi, S. R. (2014). Forensic genomics as a novel
779 tool for identifying the causes of mass mortality events. *Nature Communications*, *5*, 3652.
780 Dubey, J. P., Zarnke, R., Thomas, N. J., Wong, S. K., Van Bonn, W., Briggs, M., … Kwok, O. C. H.
781 (2003). *Toxoplasma gondii*, *Neospora caninum*, *Sarcocystis neurona*, and *Sarcocystis canis*-like
782 infections in marine mammals. *Veterinary Parasitology*, *116*(4), 275–296.
783 Duggins, D. O. (1983). Starfish predation and the creation of mosaic patterns in a kelp-dominated
784 community. *Ecology*, *64*(6), 1610–1619.
785 Eisenlord, M. E., Groner, M. L., Yoshioka, R. M., Elliott, J., Maynard, J., Fradkin, S., … van Hooidonk,
786 R. (2016). Ochre star mortality during the 2014 wasting disease epizootic: Role of population size
787 structure and temperature. *Phil. Trans. R. Soc. B*, *371*(1689), 20150212.
788 Elliott, T. A., & Gregory, T. R. (2015). What's in a genome? The C-value enigma and the evolution of
789 eukaryotic genome content. *Philosophical Transactions of the Royal Society B: Biological*
790 *Sciences*, *370*(1678), 20140331.

Fey, S. B., Siepielski, A. M., Nusslé, S., Cervantes-Yoshida, K., Hwan, J. L., Huber, E. R., … Carlson, S. M. (2015). Recent shifts in the occurrence, cause, and magnitude of animal mass mortality events. *Proceedings of the National Academy of Sciences*, *112*(4), 1083–1088.

Fuess, L. E., Eisenlord, M. E., Closek, C. J., Tracy, A. M., Mauntz, R., Gignoux-Wolfsohn, S., … Harvell, D. (2015). Up in arms: Immune and nervous system response to sea star wasting disease. *PLoS One*, *10*(7), e0133053.

Gardner, G. R., Harshbarger, J. C., Lake, J. L., Sawyer, T. K., Price, K. L., Stephenson, M. D., … Togstad, H. A. (1995). Association of prokaryotes with symptomatic appearance of withering syndrome in black abalone *Haliotis cracherodii*. *Journal of Invertebrate Pathology*, *66*(2), 111–120.

Gokudan, S., Muta, T., Tsuda, R., Koori, K., Kawahara, T., Seki, N., … Kawabata, S. (1999). Horseshoe crab acetyl group-recognizing lectins involved in innate immunity are structurally related to fibrinogen. *Proceedings of the National Academy of Sciences*, *96*(18), 10086–10091.

Goymer, P. (2007). Genetic variation: Synonymous mutations break their silence. *Nature Reviews Genetics*, *8*(2), 92.

Gravem, S. A., & Morgan, S. G. (2017). Shifts in intertidal zonation and refuge use by prey after mass mortalities of two predators. *Ecology*, *98*(4), 1006–1015.

Gudenkauf, B. M., & Hewson, I. (2015). Metatranscriptomic analysis of *Pycnopodia helianthoides* (Asteroidea) affected by sea star wasting disease. *PloS One*, *10*(5), e0128150.

Harley, C. D. G., Pankey, M. S., Wares, J. P., Grosberg, R. K., & Wonham, M. J. (2006). Color polymorphism and genetic structure in the sea star *Pisaster ochraceus*. *The Biological Bulletin*, *211*(3), 248–262.

Harvell, D., Altizer, S., Cattadori, I. M., Harrington, L., & Weil, E. (2009). Climate change and wildlife diseases: When does the host matter the most? *Ecology*, *90*(4), 912–920.

Harvell, D., Montecino-Latorre, D., Caldwell, J. M., Burt, J. M., Bosley, K., Keller, A., … Pontier, O. (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.

Hewson, I., Bistolas, K. S., Quijano Cardé, 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*, 77.

Hewson, I., Button, J. B., Gudenkauf, B. M., Miner, B., Newton, A. L., Gaydos, J. K., … Murray, M. (2014). Densovirus associated with sea-star wasting disease and mass mortality. *Proceedings of the National Academy of Sciences*, *111*(48), 17278–17283.

Hibino, T., Loza-Coll, M., Messier, C., Majeske, A. J., Cohen, A. H., Terwilliger, D. P., … Berney, K. (2006). The immune gene repertoire encoded in the purple sea urchin genome. *Developmental Biology*, *300*(1), 349–365.

Hoff, K. J. (2019). MakeHub: Fully automated generation of UCSC Genome Browser Assembly Hubs. *BioRxiv*, 550145.

Hoff, K. J., Lange, S., Lomsadze, A., Borodovsky, M., & Stanke, M. (2015). BRAKER1: Unsupervised RNA-Seq-based genome annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics*, *32*(5), 767–769.

Hoff, K. J., Lomsadze, A., Borodovsky, M., & Stanke, M. (2019). Whole-Genome Annotation with BRAKER. In *Gene Prediction* (pp. 65–95). Springer.

Hu, Z.-L., Bao, J., & Reecy, J. M. (2008). CateGOrizer: A web-based program to batch analyze gene ontology classification categories. *Online J Bioinform*, *9*, 108–112.

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2008). Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, *37*(1), 1–13.

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, *4*(1), 44.

Jackson, E. W., Pepe-Ranney, C., Debenport, S. J., Buckley, D. H., & Hewson, I. (2018). The microbial landscape of sea stars and the anatomical and interspecies variability of their microbiome. *Frontiers in Microbiology*, *9*, 1829.

Kamogashira, T., Fujimoto, C., & Yamasoba, T. (2015). Reactive oxygen species, apoptosis, and mitochondrial dysfunction in hearing loss. *BioMed Research International*, *2015*.

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., & Haussler, D. (2002). The human genome browser at UCSC. *Genome Research*, *12*(6), 996–1006.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. L. (2013). TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology*, *14*(4), R36.

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.

Lafferty, K. D., & Hofmann, E. E. (2016). *Marine disease impacts, diagnosis, forecasting, management and policy*. The Royal Society.

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*, R25.

Lasky, J. R., Des Marais, D. L., Lowry, D. B., Povolotskaya, I., McKay, J. K., Richards, J. H., … Juenger, T. E. (2014). Natural variation in abiotic stress responsive gene expression and local adaptation to climate in *Arabidopsis thaliana*. *Molecular Biology and Evolution*, *31*(9), 2283–2296.

Lessios, H. A. (2016). The Great *Diadema antillarum* Die-Off: 30 Years Later. *Annual Review of Marine Science*, *8*(1), 267–283.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*, 2078–2079.

Lieberman-Aiden, E., Van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., … Dorschner, M. O. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, *326*(5950), 289–293.

Lomsadze, A., Burns, P. D., & Borodovsky, M. (2014). Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm. *Nucleic Acids Research*, *42*(15), e119–e119.

Mah, C. L., & Blake, D. B. (2012). Global diversity and phylogeny of the Asteroidea (Echinodermata). *PloS One*, *7*(4), e35644.

Matsubara, M., Komatsu, M., Araki, T., Asakawa, S., Yokobori, S., Watanabe, K., & Wada, H. (2005). The phylogenetic status of Paxillosida (Asteroidea) based on complete mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, *36*(3), 598–605.

McCarthy, D. J., Chen, Y., & Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research*, *40*(10), 4288–4297.

Menge, B. A. (1983). Components of predation intensity in the low zone of the New England rocky intertidal region. *Oecologia*, *58*(2), 141–155.

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*(5), e0153994.

Miner, C. M., Burnaford, J. L., Ambrose, R. F., Antrim, L., Bohlmann, H., Blanchette, C. A., … Harley, C. D. (2018). Large-scale impacts of sea star wasting disease (SSWD) on intertidal sea stars and implications for recovery. *PloS One*, *13*(3), e0192870.

888 Montecino-Latorre, D., Eisenlord, M. E., Turner, M., Yoshioka, R., Harvell, D., Pattengill-Semmens, C.
889 V., … Gaydos, J. K. (2016). Devastating transboundary impacts of sea star wasting disease on
890 subtidal asteroids. *PloS One*, *11*(10), e0163190.
891 Morgulis, A., Gertz, E. M., Schäffer, A. A., & Agarwala, R. (2005). WindowMasker: Window-based
892 masker for sequenced genomes. *Bioinformatics*, *22*(2), 134–141.
893 Oberst, A., Bender, C., & Green, D. R. (2008). Living with death: The evolution of the mitochondrial
894 pathway of apoptosis in animals. *Cell Death and Differentiation*, *15*(7), 1139.
895 Paine, R. T. (1974). Intertidal community structure. *Oecologia*, *15*(2), 93–120.
896 Pancer, Z. (2000). Dynamic expression of multiple scavenger receptor cysteine-rich genes in
897 coelomocytes of the purple sea urchin. *Proceedings of the National Academy of Sciences*, *97*(24),
898 13156–13161.
899 Putnam, N. H., O'Connell, B. L., Stites, J. C., Rice, B. J., Blanchette, M., Calef, R., … Sugnet, C. W.
900 (2016). Chromosome-scale shotgun assembly using an in vitro method for long-range linkage.
901 *Genome Research*, *26*(3), 342–350.
902 Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic
903 features. *Bioinformatics*, *26*(6), 841–842.
904 Raimundo, N., Song, L., Shutt, T. E., McKay, S. E., Cotney, J., Guan, M.-X., … Shadel, G. S. (2012).
905 Mitochondrial stress engages E2F1 apoptotic signaling to cause deafness. *Cell*, *148*(4), 716–726.
906 Ralston, A., & Shaw, K. (2008). Gene expression regulates cell differentiation. *Nat Educ*, *1*(1), 127–131.
907 Ramírez-Gómez, F., & García-Arrarás, J. E. (2010). Echinoderm immunity. *Invertebrate Survival Journal*,
908 *7*(2), 211–220.
909 Saotome, K., & Komatsu, M. (2002). Chromosomes of Japanese starfishes. *Zoological Science*, *19*(10),
910 1095–1104.

911 Schiebelhut, L. M., Puritz, J. B., & Dawson, M. N. (2018). Decimation by sea star wasting disease and
912 rapid genetic change in a keystone species, *Pisaster ochraceus*. *Proceedings of the National*
913 *Academy of Sciences*, *115*(27), 7069–7074.
914 Schultz, J. A., Cloutier, R. N., & Côté, I. M. (2016). Evidence for a trophic cascade on rocky reefs
915 following sea star mass mortality in British Columbia. *PeerJ*, *4*, e1980.
916 Shadel, G. S. (2008). Expression and maintenance of mitochondrial DNA: new insights into human
917 disease pathology. *The American Journal of Pathology*, *172*(6), 1445–1456.
918 Short, F. T., Muehlstein, L. K., & Porter, D. (1987). Eelgrass wasting disease: Cause and recurrence of a
919 marine epidemic. *The Biological Bulletin*, *173*(3), 557–562.
920 Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO:
921 assessing genome assembly and annotation completeness with single-copy orthologs.
922 *Bioinformatics*, *31*(19), 3210–3212.
923 Sodergren, E., Weinstock, G. M., Davidson, E. H., Cameron, R. A., Gibbs, R. A., Angerer, R. C., …
924 Burke, R. D. (2006). The genome of the sea urchin *Strongylocentrotus purpuratus*. *Science*,
925 *314*(5801), 941–952.
926 Stanke, M., Diekhans, M., Baertsch, R., & Haussler, D. (2008). Using native and syntenically mapped
927 cDNA alignments to improve de novo gene finding. *Bioinformatics*, *24*(5), 637–644.
928 Supek, F., Bosnjak, M., Skunca, N., & Smuc, T. (2011). REVIGO summarizes and visualizes long lists of
929 gene ontology terms. *PLoS One*, *6*, e21800.
930 Tait, S. W., & Green, D. R. (2010). Mitochondria and cell death: Outer membrane permeabilization and
931 beyond. *Nature Reviews Molecular Cell Biology*, *11*(9), 621.
932 The Gene Ontology Consortium. (2016). Expansion of the Gene Ontology knowledgebase and resources.
933 *Nucleic Acids Research*, *45*(D1), D331–D338.

Tijssen, P., Pénzes, J. J., Yu, Q., Pham, H. T., & Bergoin, M. (2016). Diversity of small, single-stranded DNA viruses of invertebrates and their chaotic evolutionary past. *Journal of Invertebrate Pathology*, *140*, 83–96.

Uniprot Consortium. (2014). UniProt: A hub for protein information. *Nucleic Acids Research*, *43*(D1), D204–D212.

UniProt: The universal protein knowledgebase. (2016). *Nucleic Acids Research*, *45*(D1), D158–D169.

Wang, Y., Ning, X., Gao, P., Wu, S., Sha, M., Lv, M., … Meng, G. (2017). Inflammasome activation triggers caspase-1-mediated cleavage of cGAS to regulate responses to DNA virus infection. *Immunity*, *46*(3), 393–404.

Wares, J. P., & Schiebelhut, L. M. (2016). What doesn't kill them makes them stronger: An association between elongation factor 1-α overdominance in the sea star *Pisaster ochraceus* and "sea star wasting disease." *PeerJ*, *4*, e1876.

Wu, T. D., Reeder, J., Lawrence, M., Becker, G., & Brauer, M. J. (2016). GMAP and GSNAP for genomic sequence alignment: Enhancements to speed, accuracy, and functionality. In *Statistical Genomics* (pp. 283–334). Springer.

Yasuda, N., Hamaguchi, M., Sasaki, M., Nagai, S., Saba, M., & Nadaoka, K. (2006). Complete mitochondrial genome sequences for Crown-of-thorns starfish *Acanthaster planci* and *Acanthaster brevispinus*. *BMC Genomics*, *7*(1), 17.

Yin, T., Cook, D., & Lawrence, M. (2012). ggbio: An R package for extending the grammar of graphics for genomic data. *Genome Biology*, *13*(8), R77.

Zapico, S. C., & Ubelaker, D. H. (2013). MtDNA mutations and their role in aging, diseases and forensic sciences. *Aging and Disease*, *4*(6), 364.

Data Accessibility Statement

The *Pisaster ochraceus* nuclear genome has been deposited in the Genome database of the National Center for Biotechnology Information (NCBI; BioProject PRJNA532896, SUB5448653) and the UCSC Genome browser (http://genome.ucsc.edu); the mitochondrial genome is deposited with the NCBI (accession MH713001). RNA-Seq reads are deposited at the NCBI sequence read archive (accession no. SRP######) and the GFF3 genome annotation file and commented code are deposited in DRYAD (https://doi.org/10.6071/M3ND50).

Supplementary Information

Appendix	Supplementary methods, figures, and tables
Dataset S1	*P. ochraceus* gene model BLAST and Gene Ontology results
Dataset S2	*P. ochraceus* annotation
Dataset S3	*P. ochraceus* immune related genes
Dataset S4	*P. ochraceus* DGE by size – each size class versus all others
Dataset S5	*P. ochraceus* DGE by size – smallest versus x-largest
Dataset S6	*P. ochraceus* DGE SSWD-symptomatic versus asymptomatic
Dataset S7	Overlap between studies of *P. ochraceus* and *P. helianthoides*
Dataset S8	*P. ochraceus* RAD-seq SSWD-discriminant loci overlap with gene models
Dataset S9	Probabilities a single locus is implicated in multiple analyses
Dataset S10	Mapped read counts per library for *P. ochraceus* and *P. helianthoides* analyses
Dataset S11	Comparison of two mapping and subsampling approaches for nuclear DNA
Dataset S12	Enrichment scores and functional annotations from DAVID

Authors Contributions

M.ND. and L.M.S. conceived and designed the study with D.V.R. and J.P.W., all implemented the research; K.J.H. annotated the genome; D.V.R. and L.M.S. performed analyses, with some assistance from J.P.W. and M.ND.; D.V.R. / L.M.S., M.ND., J.P.W., and K.J.H. wrote and edited the manuscript.

The authors declare no conflict of interest.

Figures

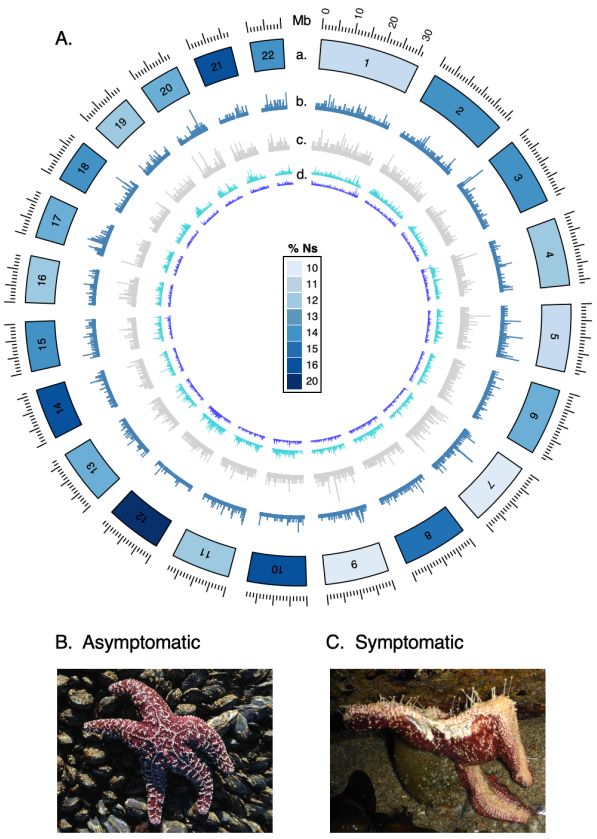

Figure 1 Genomic analysis of wasting in asymptomatic and symptomatic *Pisaster ochraceus*. *(A)* Schematic representation of the *P. ochraceus* genome. (a) Chromosome ideogram where colors represent the percentage of the super-scaffold comprised of Ns. (b) Occurrence of annotated gene models, range 1–9, and (c) occurrence of gene models without annotations, range 1–7. (d) Frequency histogram showing the proportion of 500 nt regions within contiguous 250 KB segments in which %GC content is more extreme than 2 standard deviations from the genome-wide mean: upper tail (outer track), range = 0–0.100; lower tail (inner track), range = 0–0.074 (SI Appendix, Fig. S1). *Pisaster ochraceus (B)* asymptomatic and *(C)* symptomatic for SSWD; note lesions, loss of turgor, and twisted arm, yet continued adhesion to surface with tube feet.

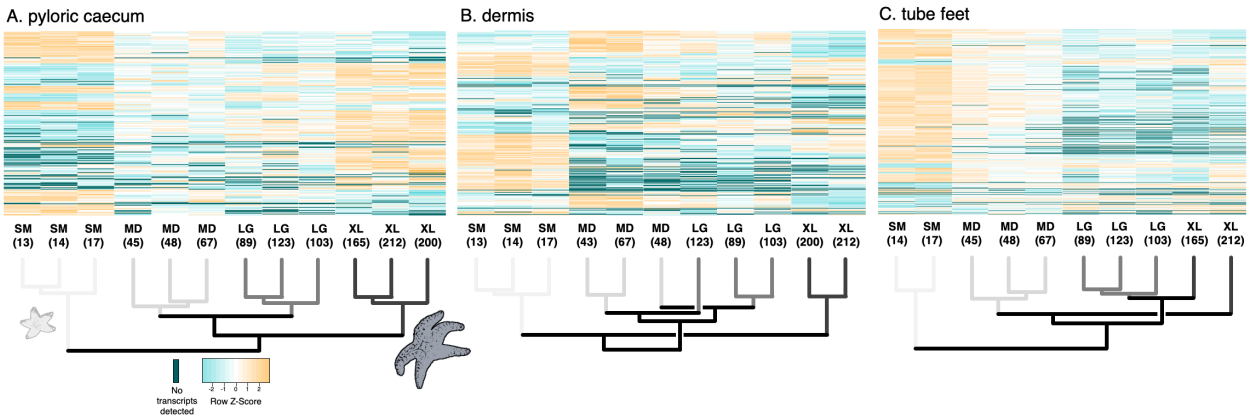

Figure 2 Differential gene expression between size-classes within tissue types. Heat maps of genes differentially expressed (FDR < 0.01) in at least one size class—small (SM), medium (MD), large (LG), and x-large (XL) (numbers represent radius in mm)—relative to the mean of all other sizes, or in SM versus XL. Heatmaps generated for each tissue type separately: *(A)* pyloric caecum, 261 genes, *(B)* dermis, 295 genes, and *(C)* tube feet, 414 genes. Colors represent log-normalized gene expression converted to Z-score. Each row corresponds to a gene, dark-teal cells correspond to NaN values. Dendrogram reflects clustering by Euclidean distance.

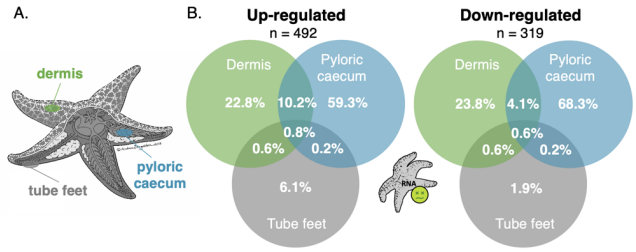

Figure 3 Summary of all differentially expressed mitochondrial and nuclear transcripts by tissue in comparisons of asymptomatic and symptomatic sea stars. *(A)* Anatomical illustration of *P. ochraceus*, highlighting tissue types used in this study. *(B)* Venn diagrams of genes up-regulated (left) and down-regulated (right) in symptomatic (relative to asymptomatic) ochre sea stars. n = number of differentially expressed genes between symptomatic and asymptomatic individuals.

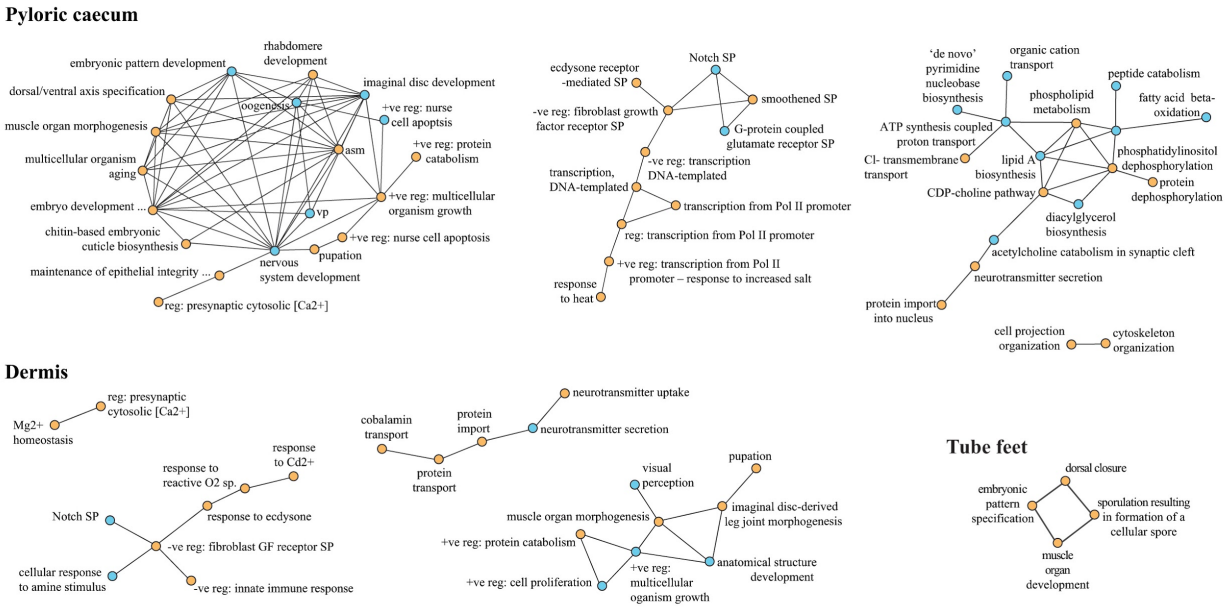

Figure 4 Associations among gene ontology (GO) terms for genes that are differentially expressed in asymptomatic versus symptomatic *Pisaster ochraceus*, organized by tissue type. Orange indicates an upregulated gene, blue a downregulated gene. Edges connect similar GO terms but do not imply co-regulation. GO terms may be specific to the organism in which the gene was identified and thus imply a related function in sea stars (e.g. pupation may represent metamorphosis). The network includes nuclear genes only, with links to mitochondrial function implied in several cases. Abbreviations: apoptosis = apoptotic process; asm = anatomical structure morphogenesis; GF = growth factor; Pol II = RNA polymerase II; SP = signaling pathway; vp = visual perception, +ve reg = positive regulation, -ve reg = negative regulation.

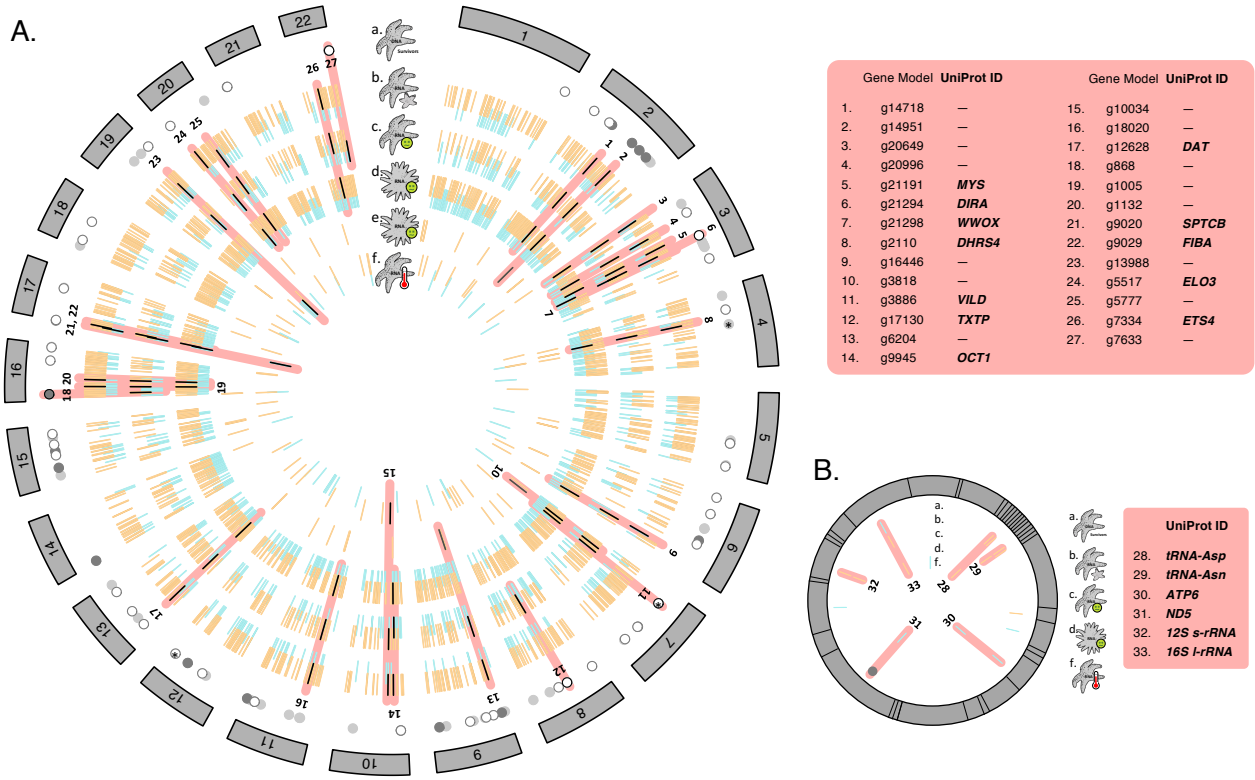

Figure 5 Comparisons of genomic and transcriptomic studies mapped to the *Pisaster ochraceus* nuclear and mitochondrial genomes. *(A)* Nuclear datasets, from the outside inwards: (a) position of the top discriminating RAD-Seq loci in *P. ochraceus* (solid = exons, open = introns, light shading = outside of gene model; n=99), * indicates the three BayeScan outlier loci identified by Schiebelhut et al. (Schiebelhut et al., 2018); (b) genes differentially expressed in smaller, relative to larger, *P. ochraceus* (FDR < 0.01), see Fig. 2 and Dataset S7a for full description of DGE by all size classes; (c) genes differentially expressed between symptomatic and asymptomatic *P. ochraceus* (FDR < 0.01); (d) genes differentially expressed between symptomatic and asymptomatic *P. helianthoides* in Fuess et al. (FDR < 0.01; Fuess et al. (2015)); (e) genes differentially expressed between symptomatic and asymptomatic *P. helianthoides* in Gudenkauf & Hewson (FDR < 0.01; Gudenkauf & Hewson (2015)); (f) genes differentially expressed between ambient ocean temperature and elevated temperature (+3°C) in *P. ochraceus* (FDR < 0.1; Chandler & Wares (2017)). *(B)* Mitochondrial datasets, from the outside inwards; icons correspond to those in panel A; DGE at FDR < 0.1. In both nuclear and mitochondrial datasets, orange = upregulated, blue = downregulated; black marks and salmon shading highlight nuclear loci recovered in ≥3 analyses (≥2 for mitochondria), or that overlapped between discriminant loci from Schiebelhut et al. (Schiebelhut et al., 2018) and track c. Numbered bars correspond to Dataset S7d.