diversity

MDPI

Article

High-Density Genetic Linkage Map of the Southern Blue-ringed Octopus (Octopodidae: *Hapalochlaena maculosa*)

Brooke L. Whitelaw [1,*], David B. Jones [1], Jarrod Guppy [1], Peter Morse [1], Jan M. Strugnell [1,2], Ira R. Cooke [3,4] and Kyall Zenger [1]

[1] Centre for Sustainable Tropical Fisheries and Aquaculture, College of Science and Engineering, James Cook University, Townsville, QLD 4811, Australia
[2] Department of Environment and Genetics, La Trobe University, Melbourne, VIC 3086, Australia
[3] College of Public Health, Medical and Vet Sciences, James Cook University, Townsville, QLD 4811, Australia
[4] Centre for Tropical Bioinformatics and Molecular Biology, James Cook University, Townsville, QLD 4811, Australia
* Correspondence: 1blwhitelaw8@gmail.com; Tel.: +61-424-642-621

Abstract: Genetic linkage maps provide a useful resource for non-model genomes and can aid in genome reassembly to form more contiguous pseudo-chromosomes. We present the first linkage map of any cephalopod, *H. maculosa*, composed of 47 linkage groups (LG). A total of 2166 single nucleotide polymorphisms and 2455 presence–absence variant loci were utilised by Lep-Map3 in linkage map construction. The map length spans 2016.62 cM with an average marker distance of 0.85 cM. Integration of the recent *H. maculosa* genome allowed 1151 scaffolds comprising 34% of the total genomic sequence to be orientated and/or placed using 1278 markers across all 47 LG. The linkage map generated provides a new perspective on HOX gene distribution in octopods. In the *H. maculosa* linkage map three (SCR, LOX4 and POST1) of six identified HOX genes (HOX1/LAB, SCR, LOX2, LOX4, LOX5, POST1) were located within the same LG (LG 9). The generation of a linkage map for *H. maculosa* has provided a valuable resource for understanding the evolution of cephalopod genomes and will provide a base for future work.

Keywords: enetic linkage; Hapalochlaena; genome evolution; cephalopods; HOX genes

Citation: Whitelaw, B.L.; Jones, D.B.; Guppy, J.; Morse, P.; Strugnell, J.M.; Cooke, I.R.; Zenger, K. High-Density Genetic Linkage Map of the Southern Blue-ringed Octopus (Octopodidae: *Hapalochlaena maculosa*). *Diversity* **2022**, *14*, 1068. https://doi.org/10.3390/d14121068

Academic Editors: Xiaodong Zheng and Ran Xu

Received: 9 October 2022
Accepted: 2 December 2022
Published: 4 December 2022

1. Introduction

Members of the blue-ringed octopus genus (*Hapalochlaena*) are the only octopods known to sequester the potent neurotoxin tetrodotoxin (TTX) within their tissues and venom [1]. They are easily identified by their iridescent blue rings and/or lines which are flashed when threatened in an aposematic advertisement of their toxicity [2]. They also remain the only octopod known to be capable of inflicting a lethal bite to humans [3,4]. Investigations into venom evolution at a transcriptomic and proteomic level have indicated lowered abundance of venom and venom-associated proteins in the venom glands of *Hapalochlaena maculosa* relative to other octopods. This is hypothesised to reflect the inclusion of TTX resulting in the redundancy of other octopod venom components [5]. Additionally, analysis of the *Hapalochlaena maculosa* genome detected a potential contraction of a key family of venom proteins (serine proteases) relative to related octopods [6].

The current genome assembly for *H. maculosa* is highly fragmented (47 K + scaffolds) [6]. Similar to other cephalopods the *H. maculosa* genome exhibits a high proportion of repetitive elements (37%), in conjunction with high heterozygosity (0.95%), and a relatively large size overall (4 Gb) [6]. These features are not conducive to generating a chromosome-level assembly [7]. The first octopod genome (the California two-spot octopus, *Octopus bimaculoides*) was published in 2015 [8] and since then an additional four octopod genomes have been published (*Callistoctopus minor, Octopus vulgaris, H. maculosa* and *Octopus sinensis*) [6,9–11]. To date, the only chromosomal-level octopus genome published is that of *O. sinensis*, [11].

Earlier karyotyping studies of cephalopods revealed a chromosomal number of 30 n for four members of Octopodiformes (*C. minor, Cistopus chinesis*, *O. vulgaris* and *Amphioctopus fangsiao*) [12,13] and 46 n for five decapodiformes (*Sepia esculenta, Sepia lycidas, Sepioteuthis lessoniana*, *Heterololigo bleekeri* and *Photololigo edulis*) [12]. While genetic resources such as genomes are gaining traction in the elucidation of cephalopod evolution, they are difficult to generate and do not provide a complete picture.

Linkage maps are a valuable tool that facilitates the examination of key processes governing genome evolution by acting as a platform for analysing large (chromosomal rearrangements) and fine-scale (gene-specific) events [14]. They form an important component of comparative whole genomic analyses, providing support for genome assembly and annotation in non-model species [15]. Historically, genetic linkage maps have been used to elucidate the inheritance of complex disease traits and for marker-assisted selection in agricultural species via Quantitative Trait Loci (QTL) analyses [16]. More recently linkage maps have been utilised in non-model species in an ecological context, aiding the untangling of genotypic and phenotypic traits and their modes of selection [17,18].

Genetic linkage maps have great utility when applied to cephalopod genomics and could improve our understanding in three key ways, (i) improvement of currently fragmented genome assemblies, (ii) supporting the assembly of new genomes, improving ease and reliability and (iii) bridging the gap between genetics and phenotype. Chromosome-level genome assemblies are also valuable because they allow for the examination of the role of large-scale structural variation in evolutionary processes such as speciation, while also providing the genomic context to understand the evolution of key gene families. The HOX gene cluster is a highly conserved set of developmental genes occurring throughout metazoans [19]. Cephalopods exhibit drastic changes to their body plan relative to other metazoans courtesy of unique expression patterns of recruited HOX genes [20]. A total of nine genes were first identified in the Hawaiian bobtail squid *Euprymna scolopes* [21] and later found to be shared with *O. bimaculoides*. However, unlike related molluscan genomes [22] the HOX genes did not form a cluster and were instead scattered across different scaffolds ranging in length between 53 kb and 751 kb [8]. Although this suggests that there is no HOX cluster in cephalopods, the potential co-location of cephalopod HOX genes at the chromosome level has not yet been investigated. Furthermore, genetic linkage maps for one species can aid in the generation of maps for related taxa as has been demonstrated for the reed warbler (*Acrocephalus arundinaceus*) whose linkage map was derived from the common chicken [23]. Lastly, a genetic linkage map also provides a method to link octopod genomics and phenomics, facilitating the investigation of the genetic basis of biological traits. Such studies have applications in evolutionary biology (QTL associated with retention of paedomorphic traits in salamanders) [24] and aquaculture (body weight and length of turbot) [25]. QTL studies could be valuable for commercial species such as *Octopus vulgaris* which are within the early stages of development as an aquaculture species [26].

Generation of a linkage map for cephalopods has not previously been attempted, in part due to the absence of genome assemblies and genetic marker resources prior to 2015 [8], but also due to the difficulty in keeping and breeding cephalopods in captivity. This latter requirement is essential to track pedigree information and attain the large family sizes, which are required for robust linkage analysis [27]. This study presents the first linkage map of any cephalopod, the Southern Blue-Ringed Octopus (*Hapalochlaena maculosa*) constructed using a combination of single nucleotide polymorphism (SNP) and presence–absence variant (PAV) markers.

2. Materials and Methods

2.1. Sample Collection and Family Structure

Detailed descriptions of sample collection, mating structure and housing are available in Morse et al., 2018 [28]. Briefly, specimens of *H. maculosa* were collected from Cockburn Sound, Western Australia using false shelter traps. Animals were individually housed

at the Fremantle Octopus facility in 1 litre flow-through aquaria within a 1000 L sump. Collected animals (36 in total) were divided into 12 groups with each female being paired with one of the two males sequentially. Animals were paired in a larger separate 30 L aquarium overnight and interactions recorded. Resulting families' clutch sizes are located in Table S1.

2.2. DNA Extraction and SNP Generation

Muscle tissue samples were taken from the arm tip of adults and the body of embryos. DNA was extracted from samples using the CTAB protocol followed by purification with a Sephadex G-50 (GE Healthcare Life Sciences 2000). Quality and quantity of genomic DNA were assessed through visualization on a 0.8% agarose gel. SNP discovery was conducted by Diversity Arrays Technology PL, Canberra ACT, Australia using a restriction digest method, specifically enzymes *Pstl* and *Hpall* in conjunction with proprietary barcodes. Samples exhibiting uneven digestion patterns were excluded from library preparation. Pooled libraries were sequenced using a HiSeq2500 and resulting reads were filtered for quality (Q < 25) [29]. SNPs were called on the subsequent data set using the KDcompute pipeline (DArT) [30].

2.3. SNP Selection, Quality Control and Genotyping

Filters were applied to the SNP (Single Nucleotide Polymorphisms) data provided by Diversity Arrays using the program DARTQC (https://github.com/esteinig/dartQC accessed on 5 June 2017). Retained SNPs met the following criteria, individual genotypes were called on a minimum of 5 read counts, average SNP repeatability > 0.9, SNP call rate > 0.6 and MAF (Minor Allele Frequency) > 0.01. In addition, sequence similarity searches of 95% identity were conducted within CD-HIT [31,32] and for any SNPs that were identified within the same flanking regions, only the SNP with the highest MAF was retained. Individual families were identified and PLINK [33] was used to identify and remove mendelian errors at a threshold of 0.1 with any remaining errors set to missing. Presence–absence variant (PAV) loci were filtered for reproducibility >0.9 and a chi-square test performed to identify mendelian errors ($p > 0.05$). SNP markers were merged across ten families using bedtools merge [34], SNPs were included if they occurred in a single family. PAV and SNP markers for ten families were combined to form a single file, and only markers that occurred in at least one family were retained.

2.4. Paternity and Family Designation

Due to the potential of copulations prior to capture and the ability of females to produce offspring fathered by multiple males [28], paternity was evaluated using 7100 of the filtered SNPs. Cervus v3.0.7 [35] was used to calculate allele frequencies across all samples, simulate parentage and assign parent to offspring parameters used were as follows (parent pairs with known sexes, 1600 offspring, 15 candidate mothers, probability of mother sampled 0.95, 20 candidate fathers, probability of father sampled 0.95, probability of loci typed 0.95, probability of loci mistyped 0.01 and minimum loci typed 200). Colony was also used to validate families with the following parameter settings (female monogamy, male polygamy, no inbreeding, no clones, diecious, diploid, full-likelihood, medium precision, updating allele frequencies—yes, sibship scaling—yes, weak prior, paternal sibship size 20, maternal sibship size 20, 968 markers, 511 progeny and respective parents) [36].

2.5. Map Construction

The Lep-MAP3 pipeline [37] was used to construct the final map composed of 10 families with a total of 277 individuals. Markers were filtered for high levels of segregation distortion (parameter dataTolerance 0.001, X2 test, df = 1–2, $p < 0.001$) using the Filtering2 module. Missing parental genotypes were then predicted using the ParentCall2 module with half siblings taken into account with the–halfsib flag. SeparateChromosomes2

was used to assign markers to linkage groups (LG) based on a LOD threshold of 10 and minimum group size of 10.

The resulting 47 LG were ordered using OrderMarkers to create a sex-averaged map. Sex informative maps were generated using the preserved marker order obtained from the sex-averaged map with intervals recalculated based on female and male informative meiotic events. Maps were visualised using the LinkageMapView package in R.

2.6. Re-Orientation of Genomic Scaffolds and Genes of Interest

In order to improve contiguity of the *H. maculosa* genome, markers were mapped to their genomic locations and used in conjunction with the generated linkage map to re-orientate and/or place scaffolds. This produced 47 pseudo-chromosomes, each corresponding to a linkage group. SNP and PAV markers were mapped to the genome of the Southern Blue-Ringed Octopus (NCBI: PRJNA602771, *Hapalochlaena maculosa*) [6] using the bwa mem [38] with a seed length of k = 20. Input files for chromonomer were generated using the output from Lep-Map3 in R in conjunction with the mapped loci locations. Chromonomer was run with default parameters [39] including the following inputs: genomic locations for mapped markers and contigs, genome.fasta and genetic linkage map structure. To determine the proportion of the genome included in the 47 pseudo-chromosomes the sum of base pairs from each pseudo-chromosome generated by chromonomer was divided by the total genome length (bp). Assembly stats (https://github.com/sanger-pathogens/assembly-stats accessed on 15 October 2019) was run to assess the new pseudo-chromosome assembly produced by chromonomer.

Conserved HOX genes were identified from the *H. maculosa* genome by conducting BLAST searches against a database of known cephalopod HOX genes manually curated from NCBI. The mutual best hits were retained and annotated using InterPro [40]. Genomic positions were ascertained by aligning genes to the genome using est2genome from exonerate v2.2 [41]. Cephalopod HOX genes were aligned using MAFFT v7.407 [42] and a consensus maximum likelihood phylogeny generated using RAxML v 8.0 with the WAG substitution model for 100 bootstraps [43].

2.7. Genome Coverage

Genome coverage was ascertained by estimated genome size/summed length of all linkage groups. Genome size was estimated in centimorgans (cM) using two methods [44] and the average taken for coverage estimation. It should be noted that genome size here does not refer to a physical distance but to a distance calculated using recombination frequencies similar to the linkage map. The average interval for each linkage group (LG) was calculated by dividing total LG length by the number of intervals. Genome size estimation 1 (Ge 1) multiplied by the total LG length by the factor $(m + 1)/(m - 1)$, whereby m = number of markers within each LG [45]. Genome size estimation 2 (Ge 2) was calculated by adding 2*(average marker interval) to the total LG length [46]. The sum across all LGs was calculated with each method and the average between the two used.

2.8. Segregation Distortion and Sex-Specific Recombination

Gametic and post-zygotic selection can result in deviations from Mendelian inheritance of codominant alleles known as segregation distortion. In order to detect segregation distortion between families and across markers log-likelihood ratio tests were performed for goodness of fit to Mendelian expectations. Tests were performed using the software LINKMFEX v3.4 [47] which calculates G-values across all parents for each family. Insufficient number of loci significantly deviated from expected segregation ($p > 0.05$) within and between families to perform further testing with confidence (Table S6). Sex-specific recombination rates were ascertained using LINKMFEX v3.4. Intervals shared between genders independent of family were identified in R and are located in Table S5. Due to the low number of intervals shared between sexes, sex-specific recombination rates could not be calculated nor compared with confidence.

3. Results

3.1. Genotyping and Paternity

A total of 25,253 presence–absence variants (PAV) loci and 19,729 SNPs were discovered using DArTSeq™ genotyping (Diversity Arrays) [28]. Stringent filtering, as previously described (methods Section 2.3), resulted in the retention of 5656 SNPs and 11,179 PAV for linkage mapping analysis (Tables S2 and S3). Parentage analysis revealed multiple paternity in 12 clutches resulting in 28 families and 543 offspring in total. Additionally, of the twenty-eight families, pre-capture copulation of five females resulted in five families fathered by a non-genotyped male, and as such, these were excluded from downstream analyses. Families with fewer than 10 individuals or with a parent that did not pass PLINK filtering thresholds (Mendelian errors > 10% per marker and 10% per individual) were also excluded. The remaining ten families (seven females and ten males) were used in linkage map generation (Table S1).

3.2. Linkage Map Construction and Genome Coverage

A high-density sex-averaged map was generated using 4621 informative loci across 10 families and resulted in 47 linkage groups (LGs) (Table 1, Figure 1). The total map length was 2016.62 cM with LGs ranging between 4.97 cM (LG 36) and 113.68 cM (LG 7) and averaging to 42.9 cM. Intervals between markers averaged to 0.85 cm (0.44 cM, 0 cM intervals inclusive) with the estimated chromosome number for *H. maculosa*, based on other octopods being 30 [13]. Markers per linkage group ranged between 370 (LG 1) to 10 (LG 47), with 20 LGs containing >80 markers. Sex-specific maps were relatively shorter with maternal and paternal maps totalling 1860.89 cM and 1832.08 cM in length, respectively. Intervals between markers were slightly greater in the maternal map compared to the paternal map. Paternal maps exhibited an average interval of 1.38 cM (0.53, 0 cM intervals inclusive) while maternal maps were on average 1.46 cM (0.51, 0 cM intervals inclusive). More markers were found to be informative in the maternal (3681) map relative to the paternal (3478). Overall, linkage group length ranged from <1 cM (LG 44 and 45) to 102.5 cM (LG 2), and from <1 cM (LG 43)–105.93 (LG 1) for maternal and paternal maps, respectively. Linkage group lengths between maternal and paternal maps were highly comparable. Genome coverage for sex average, maternal and paternal maps was estimated to be 96.4%, 94% and 97.2%, respectively (Table 2).

Table 1. Statistical summary of *Hapalochlaena maculosa* sex averaged, paternal and maternal linkage maps. Linkage groups (LG).

LG	Total Length (cM)			Number of Markers			Average Interval (cM)		
	Sex Average	Female	Male	Sex Average	Female	Male	Sex Average	Female	Male
1	101.31	88.98	105.93	370	305	279	0.27	0.29	0.38
2	102.98	102.05	91.81	356	289	269	0.29	0.35	0.34
3	95.24	84.20	95.51	354	288	272	0.27	0.29	0.35
4	63.49	50.57	62.69	287	244	215	0.22	0.21	0.29
5	93.96	90.51	96.63	257	210	187	0.37	0.43	0.52
6	86.02	66.12	68.29	254	189	198	0.34	0.35	0.35
7	113.68	89.83	98.66	243	196	189	0.47	0.46	0.52
8	55.50	34.80	50.94	234	150	194	0.24	0.23	0.26
9	78.21	82.11	82.55	208	149	165	0.38	0.55	0.50
10	44.96	43.07	40.95	184	154	129	0.25	0.28	0.32
11	62.05	50.22	70.16	163	136	105	0.38	0.37	0.67
12	66.77	64.19	61.34	161	120	126	0.42	0.54	0.49

Table 1. *Cont.*

LG	Total Length (cM)			Number of Markers			Average Interval (cM)		
	Sex Average	Female	Male	Sex Average	Female	Male	Sex Average	Female	Male
13	53.70	52.67	55.69	138	111	117	0.39	0.48	0.48
14	65.09	37.14	67.88	126	115	81	0.52	0.33	0.85
15	58.09	59.86	50.87	125	98	92	0.47	0.62	0.56
16	40.05	42.89	35.54	125	99	92	0.32	0.44	0.39
17	60.77	56.83	77.62	113	93	79	0.54	0.62	1.00
18	54.33	51.58	54.52	94	74	74	0.58	0.71	0.75
19	59.61	63.96	50.76	91	71	73	0.66	0.91	0.70
20	54.87	46.72	45.57	82	65	63	0.68	0.73	0.74
21	56.85	48.76	48.77	73	57	51	0.79	0.87	0.98
22	36.55	30.84	37.58	66	52	51	0.56	0.60	0.75
23	27.91	30.26	25.38	49	37	40	0.58	0.84	0.65
24	30.36	29.02	17.11	43	38	25	0.72	0.78	0.71
25	46.10	57.29	40.64	38	26	36	1.25	2.29	1.16
26	33.59	31.70	16.67	35	34	16	0.99	0.96	1.11
27	33.13	32.50	32.35	34	28	26	1.00	1.20	1.29
28	27.44	32.76	11.96	28	22	22	1.02	1.56	0.57
29	33.12	41.40	22.82	28	24	17	1.23	1.80	1.43
30	25.93	16.98	21.63	27	20	18	1.00	0.89	1.27
31	16.90	20.27	13.62	26	23	22	0.68	0.92	0.65
32	26.47	24.32	32.99	23	19	21	1.20	1.35	1.65
33	21.44	36.84	17.30	20	17	15	1.13	2.30	1.24
34	26.78	41.37	12.59	18	16	15	1.58	2.76	0.90
35	24.86	13.57	42.30	16	12	12	1.66	1.23	3.85
36	4.97	2.13	1.32	14	13	8	0.38	0.18	0.19
37	16.39	6.67	11.58	13	8	10	1.37	0.95	1.29
38	8.71	3.34	6.38	12	9	10	0.79	0.42	0.71
39	6.74	6.91	4.31	11	5	10	0.67	1.73	0.48
40	19.26	23.48	6.94	11	10	8	1.93	2.61	0.99
41	15.74	21.25	8.01	11	9	7	1.57	2.66	1.33
42	16.22	12.45	3.85	10	8	7	1.80	1.78	0.64
43	14.57	21.06	0	10	10	3	1.62	2.34	0.00
44	5.79	0	8.41	10	2	10	0.64	0.00	0.93
45	7.81	0	12.28	10	7	9	0.87	0.00	1.53
46	10.37	10.17	6.91	10	10	3	1.15	1.13	3.45
47	11.96	7.24	4.55	10	9	7	1.33	0.91	0.76

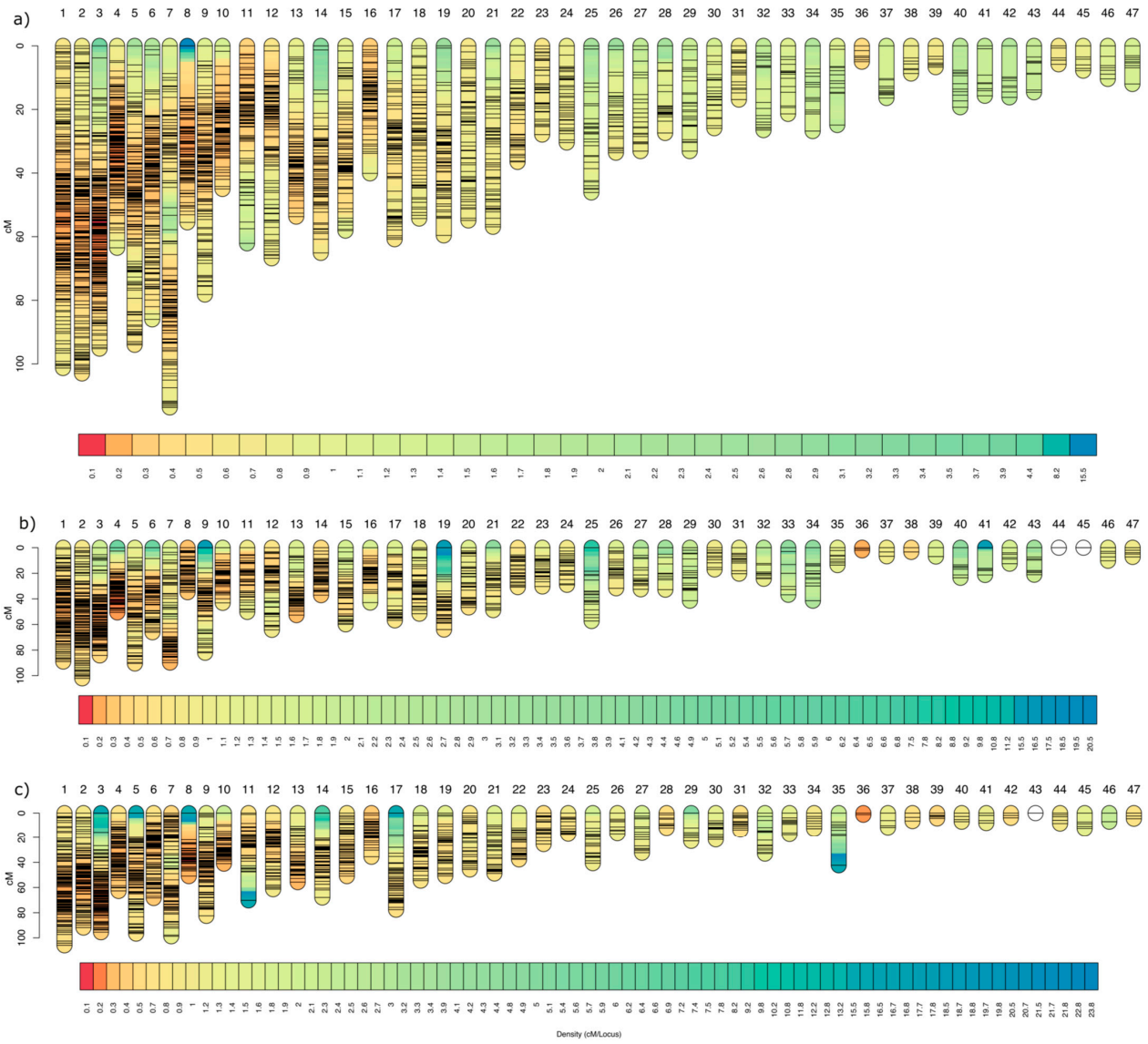

Figure 1. Genetic linkage map generated using LEPMAP3 for *H. maculosa* (**a**) sex-averaged map (**b**) maternal map (**c**) paternal map. Marker density (cM/Locus) is visualised in a scale from red (high) to blue (low).

Table 2. Genome coverage estimation of sex average, maternal and paternal *Hapalochlaena maculosa* linkage maps. Genome length estimation method 1 (Ge1) and 2 (Ge2).

Method	Map	Ge Result	Average between Methods Ge1 + Ge2/2	% Coverage
Ge1 = LG length * (marker number + 1/marker number − 1)	Sex average	2091.75	2091.75	96.41
Ge2 = LG length + (2 * average interval)	Sex average	2091.75	2091.75	
Ge1 = LG length * (marker number + 1/marker number − 1)	Male	1914.07	1914.06	97.22
Ge2 = LG length + (2 * average interval)	Male	1914.06	1914.06	
Ge1 = LG length * (marker number + 1/marker number − 1)	Female	1949.44	1949.44	93.98
Ge2 = LG length + (2 * average interval)	Female	1949.44	1949.44	

3.3. Genome Mapping and Scaffold Reorientation

Consolidation of genetic linkage and sequence maps was achieved using 1278 markers across 47 LGs which successfully arranged 1151 scaffolds covering 34% (1.39 GB) of the sequenced genome for *H. maculosa* (Table 3). Orientation was possible in 105 scaffolds with >1 SNPs mapped, the remaining 1239 scaffolds with a single mapped marker were able to be placed within a linkage group. Splits occurred in 10 genomic scaffolds to maintain congruence of linkage map marker order, which was given precedence over genomic contig order, additionally, 286 inconsistent markers were excluded.

Table 3. Assembly statistics for the original *Hapalochlaena maculosa* assembly and the new chromonomer-generated assembly.

Metrics	Original Genome Assembly	Chromonomer Assembly
Total length (Mb)	4009.60	4009.63
Number of Scaffolds	48,284.00	47,190.00
Largest scaffold (Mb)	11.01	150.55
Average scaffold length (Mb)	0.08	0.08
N50 (Mb)	0.93	1.25
N60 (Mb)	0.65	0.79
N70 (Mb)	0.44	0.49
N80 (Mb)	0.27	0.28
N90 (Mb)	0.12	0.13
N100 (Mb)	0.01	0.01
N count (Mb)	1574	1574
Gaps (Mb)	4.32	4.32

Pseudo-chromosome lengths varied greatly with the largest scaffold LG 3 (150 Mb) dwarfing the smallest LG39 (6049 bp), with an average length of 29.58 Mb. Promising scaffolds containing a minimum of one mapped marker were absent from the linkage map and therefore excluded (n = 311). Conflicting markers were identified in 203 scaffolds indicating possible assembly errors within the genome. A total of 5876 genes were annotated across all pseudo-chromosomes.

The most common Pfam domain within the gene set was cadherins, which were present in 53 genes (Table S8). The new assembly resulted in an overall reduction in the number of scaffolds from 48 K to 47 K and the largest scaffold increased from 11 Mb in the original assembly to 150 Mb. A summary of chromonomer results is available at this link (http://203.101.230.130/chromonomer/index.php?v=hmac2020 accessed on 23 October 2019).

3.4. Genes of Interest (HOX)

HOX genes, which are integral to development, and are conserved throughout metazoans, were identified from the annotated *H. maculosa* genome and located within the updated assembly. Of the eight HOX genes identified within octopods, six were present in *H. maculosa* assembly (HOX1/LAB, SCR, LOX2, LOX4, LOX5 and POST1) (Figure 2). All HOX genes occurred on separate scaffolds within the original assembly, however the updated assembly was able to place all three genes within LG 9 (SCR, LOX4 and POST1). The remaining three genes (HOX1/LAB, LOX2 and LOX5) were not present within a linkage group and were located on separate scaffolds (Figure 2). A total of four HOX genes (HOX1/LAB, LOX2, LOX5 and POST1) were also identified from the chromosomal level genome assembly of *O. sinensis* [11]. HOX genes identified in *O. sinensis* could not corroborate the arrangement observed in *H. maculosa* LG9 as only one of the three genes (POST1)

was identified. Phylogenetic analysis of HOX genes identified from six cephalopod species, three octopods (*O. bimaculoides*, *O. sinensis* and *H. maculosa*), two bobtail squids (*Euprymna scolopes* and *Euprymna tasmanica*) and a cuttlefish (*Sepia officinalis*) displayed high bootstrap support (>70) [48] for all the gene clades observed (Figure 3) in accordance with findings by Hills and Bull, 1993.

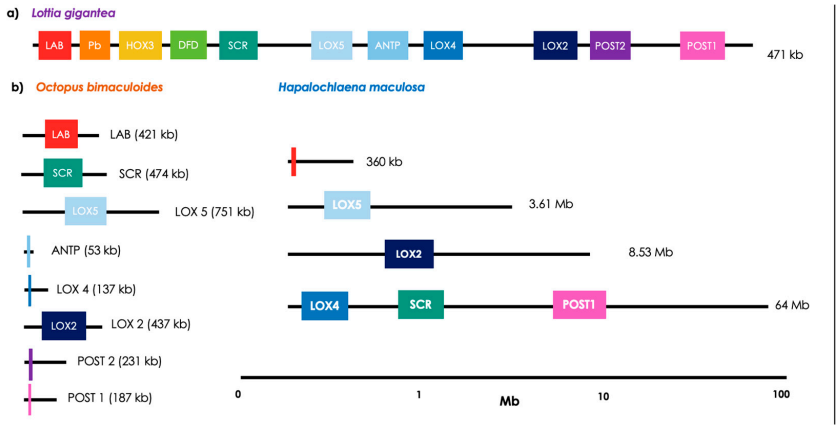

Figure 2. Comparison of HOX gene arrangement in the (**a**) owl limpet (*Lottia gigantea*), (**b**) California Two-Spot Octopus (*Octopus bimaculoides*) and The Southern Blue-Ringed Octopus (*Hapalochlaena maculosa*). Genes are coloured consistently between species. Scaffold lengths for (**b**) follow a logarithmic scale.

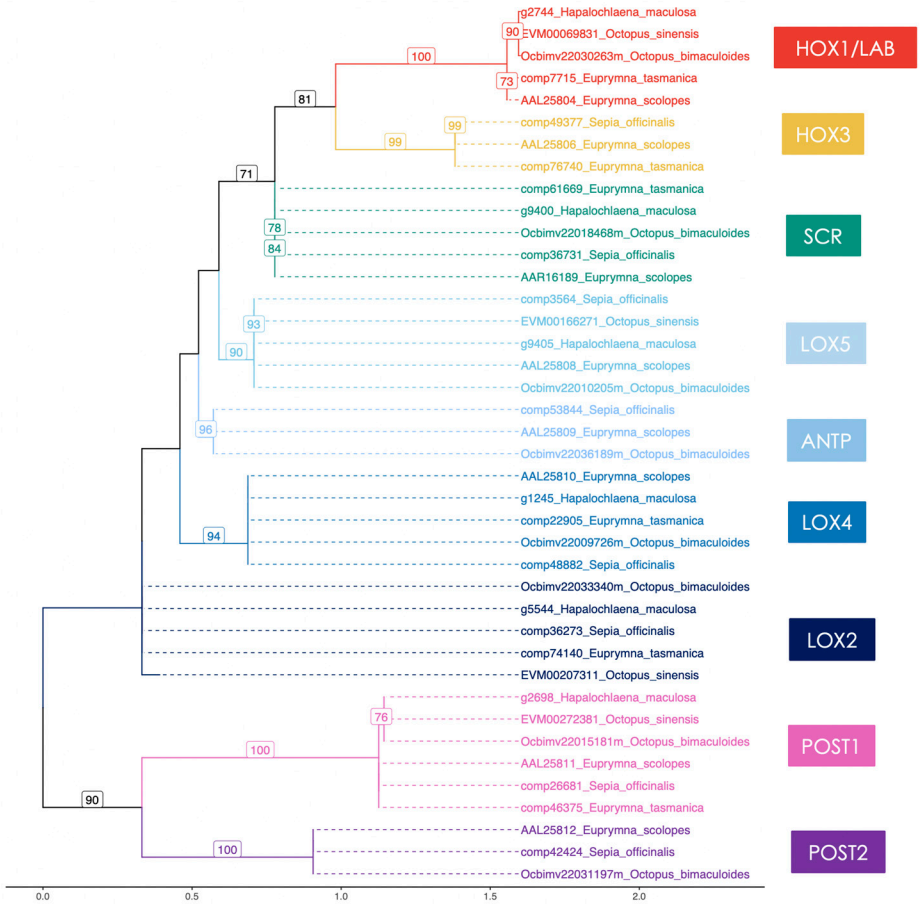

Figure 3. Maximum likelihood phylogeny of the HOX genes (aa) in six cephalopods. Bootstrap values >70 present on branches. Taxa are coloured in accordance with HOX genes.

4. Discussion

This study presents a high-density genetic linkage map for the Southern Blue-Ringed Octopus (*Hapalochlaena maculosa*), the first generated for any cephalopod. Integration of the genetic linkage maps with the current genome assembly [6] facilitated the consolidation and reorientation of scaffolds into a more contiguous assembly. The improved assembly provided new insights into the genome evolution of cephalopods, revealing the genomic placement of several key developmental HOX genes, which has previously not been possible in any other cephalopod genome assembly to date. Genetic resources generated in this study pave the way for future studies examining genomic and functional evolution across cephalopods.

4.1. Paternity and Linkage Map Generation

Robust linkage map generation is highly dependent on the number of families, adequate family sizes and reliable pedigree reconstruction [49]. Larger families present greater opportunities to detect informative meiotic events, which are crucial in linkage analysis. Linkage maps are generated by calculating recombination events between loci and are thus dependent on the number of recombination events observed for the statistical strength of these calculations. The relatively smaller family sizes present in this study result in fewer recombination events observed between loci. This demonstrates some of the challenges associated with using wild-caught octopods for linkage map generation. Even though mapping can be conducted across small families, it is estimated that a minimum of 200 individuals was ideal for map generation in several simulated data sets (recombinant inbred lines, F2 with dominant and co-dominant markers, double haploid and backcrossing) [50]. However, standardization of linkage map generation is difficult due to the differences in genome structure (i.e., size and number of chromosomes), recombination rates and variation of life history traits between organisms which impacts the size and structure of families available to be sampled [27]. The data presented here provided the best genetic linkage map available to date in cephalopods, despite the small relative clutch sizes.

This study utilised data generated from a complimentary multiple paternity study on *H. maculosa* with parental samples sourced from the wild (Cockburn Sound, Western Australia) [28], while this was the most successful captive breeding attempt to date in this genus and rendered our current work possible, the data were not without limitations for the purposes of the current study. Before capture, and during the multiple paternity study, females were exposed to multiple males resulting in observed cases of multiple paternity. All females exhibited multiple paternity, including five unknown males (presumably from mating in the wild prior to the females capture). As a result, family sizes were reduced as female clutches were divided among the multiple males, which resulted in the loss of nine families, each containing < 10 individuals, in addition to the five families containing unknown paternity. To address these challenges, a linkage mapping algorithm was selected that was able to use a combined dataset integrating all families to maximise power in linkage group determination and marker order. Lep-Map3 was employed in this study because it allows not only for the integration of all families as one dataset, but also takes into account sibship between offspring [37].

A total of 47 linkage groups were resolved using 4621 markers with a sex average map length of 2016 cM (average interval length of 0.85). Map length in *H. maculosa* is larger than many other published molluscan linkage maps including of the bivalve *Pinctada maxima* (831.7 cM) [51] and the gastropod *Biomphalaria glabrata* (746.7 cM) [52], which have relatively smaller genomes 290 Mb (*Neomenia permagna*)–2.21 Gb (*Bathymodiolus plantifrons*) [53] compared to *H. maculosa* (4 Gb). This suggests a similar recombination rate between molluscan lineages. Modelling conducted by Stapley et al. (2017) demonstrated a negative relationship between genome size and recombination rate within plant genomes. However, animals and fungi did not exhibit a reduction in recombination rate with an increase in genome size [54]. Genomic features may also impact recombination rates with the density of repetitive elements correlated negatively, while gene density exhibits a

positive relationship with recombination rate [55–58]. Fine-scale analysis of recombination rates along linkage groups was not possible with the current linkage map of *H. maculosa* generated in this study. However, future linkage maps could be utilized to examine patterns of recombination within the large and repeat-rich genomes characteristic of cephalopods.

4.2. Cephalopod Genome Evolution and Structure

Prior to this study, no genetic linkage maps have been produced for any cephalopod; however, karyological studies conducted on a subset of 11 cephalopods estimate chromosome numbers for octopodiformes (n30), decapodiformes (n46) and nautiloids (n26) [12,13,59]. Karyograms revealed evolutionary distances between species to be congruent with molecular phylogenies conducted on the three octopods, *C. minor, A. fangsiao* and *C. chinensis* [13]. Linkage groups generated in this study do not correspond 1:1 with the expected chromosome number for *H. maculosa* suggesting the need for either a larger number of families, a larger number of progenies per family, or a larger set of informative markers to allow for chromosome level resolution and improved contiguity.

Cephalopods are characterised by their large and repetitive heterozygous genomes ranging between 3 Gb (*O. bimaculoides*) and 5 Gb (*C. minor*) with 46% and 44% repetitive element composition [8,9]. These characteristics are compounded by the use of short-read sequencing techniques to generate many older genomes often resulting in difficult-to-assemble and highly fragmented genomes [53]. Long-read sequencing technologies such as Pac-Bio and Oxford Nanopore have improved the quality and contiguity of genomes, however, factors such as cost and DNA quality can be prohibitive in some cases [60]. The *H. maculosa* genome published recently was highly fragmented and was composed of 48 K scaffolds [6]. After integration with the linkage map generated in this study, the genome assembly was reorganized into 47 pseudo-chromosomes covering 1.3 Gb and integrating 34.6% of the total genome. It should be noted that pseudo-chromosomes as used here denote the closest to a chromosome assembly currently produced for a member of this genus. This improved genome assembly could be applied to examine evolutionary history of genomes, in addition to assisting in the generation of future octopod genomes [61]. Fragmented non-model genomes greatly benefit from complementary linkage maps, which allow for the reorientation and placement of scaffolds into a chromosome-level assembly [62]. Because linkage maps are independent of the primary assembly they can be used for the continuous improvement of genomes as new primary assemblies become available [63]. This issue is highlighted by the results of this study where our highly fragmented primary assembly for *H. maculosa* meant that comparatively few scaffolds contained SNPs that could be incorporated into the linkage map. Future improvements in the primary assembly (e.g., by long-read sequencing) could make use of the linkage map presented here to achieve a much more contiguous final result.

4.3. Evolution of the HOX Gene Cluster in Cephalopods

Due to the improved contiguity of the *H. maculosa* genome, six of the eight expected genes from the HOX cluster were placed into a greater genomic context than has previously been possible in cephalopods. The HOX gene cluster forms a highly conserved set of developmental genes in metazoans [19]. A single set of nine genes was first observed in *E. scolopes*, and corroborated in the *O. bimaculoides* genome as a coleoid-specific trait [8]. HOX genes in *E. scolopes* were assumed to form a cluster based on the pattern of expression along the central nervous system (CNS), which was found to be congruent with the ancestral role of axial patterning [20]. Examination of homologous genes in *O. bimaculoides* revealed rearrangements of HOX genes inconsistent with other classes present within Mollusca as each gene was located on a separate scaffold [8]. However, greater resolution of each gene's placement in relation to each other was not possible due to the fragmentation of the genome [8]. The improved assembly for *H. maculosa* generated in this study was able to place three HOX genes within the same scaffold (LOX4, SCR and POST1). Genomic placement of HOX genes within the chromosome-level assembly of *O. sinensis* could not

support or contradict the co-location of genes within *H. maculosa* as only a single gene (POST1) could be identified. Implications behind HOX gene placement in *H. maculosa* are difficult to infer without expression studies. Unfortunately, patterns of HOX gene expression have not yet been examined in octopods and the implications of these gene placement rearrangements in *H. maculosa* remain unknown.

4.4. QTL Mapping and Future Work

Linkage maps provide a framework for understanding the evolution and inheritance of particular phenotypic traits through qualitative trait loci (QTL) studies [64]. Historically, such studies have been used to improve desirable traits in agricultural and aquacultural species including disease resistance [65], fibre quality (cotton) [66] and growth rate (Asian seabass) [67], to name a few. Aquacultural significance may not be applicable to *Hapalochlaena* directly, however, an increasing number of octopod species are being raised for human consumption with 745,054 tonnes produced in 2003 within Japan alone [68]. *Octopus vulgaris* is a prime target for aquacultural production due to their relatively large clutch sizes and growth rates [69]. The *H. maculosa* linkage map produced in this study may aid in the prediction and construction of similar maps for related octopods, in addition to supporting the construction of additional octopod genomes. A similar application has been conducted when the draft chicken genome was used to predict the linkage map of passerine birds, which was then verified using the more closely related reed warbler (*Acrocephalus arundinaceus*) linkage map [23]. Despite the divergence times between the Galliformes (chicken) and Passeriformes lineages (~80–100 mya) [70,71] sufficient synteny and microsatellite markers were conserved to facilitate map linkage map construction [23]. Furthermore, linkage maps and resultant QTL studies have applications in unravelling the evolution of complex multigenic traits associated with adaptation and speciation [72,73]. A link was observed between QTL associated with male song and female song preference for two cricket species *Laupala kohalensis* and *Laupala paranigra* [64]. While the genes underlying the traits have yet to be identified, analysis of QTL allows for the investigation of evolutionarily relevant traits and provides a guide for future genomic work [64].

5. Conclusions

This study successfully fulfilled the primary aim by producing the first linkage map of a cephalopod, the Southern Blue-Ringed Octopus (*Hapalochlaena maculosa*). Integration of the linkage map with the current genome assembly reduced fragmentation and enabled the placement of three HOX genes within the same linkage group providing a greater genomic context for this gene family compared to currently available cephalopod genomes. The linkage map produced in this study will provide a valuable resource for the generation of *Hapalochlaena* genomes while also providing a framework for understanding the inheritance of phenotypic traits through future work.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/d14121068/s1, Genotype_calls_DArTseq_Hmaculosa.xlxs: Genotype data for all *H. maculosa* families; Table S1: Summary of *Hapalochlaena maculosa* families; Table S2: Filtering and quality assessment of SNPs produced using DARTQC; Table S3: SNPs removed through filtering in ten *Hapalochlaena maculosa* families using PLINK; Table S4: Characteristics and summary of *Hapalochlaena maculosa* Chromonomer assembly; Table S5: Recombination rates for shared loci intervals in maternal and paternal maps; Table S6: Significant G-values for segregation distortion test performed using LINKMFEX; Table S7: HOX gene summary; Table S8: The top 30 Pfam domains for genes mapped within the 47 pseudo-chromosomes generated using chromonomer; Table S9: Genome coverage estimation of sex average, maternal and paternal linkage maps.

Author Contributions: B.L.W.: Project development, computational analyses, results interpretation, figure generation, primary manuscript author. D.B.J.: Bioinformatics support, results interpretation and aided with editing/proofing. J.G.: Bioinformatics support, results interpretation and aided with editing/proofing. P.M.: Conducted an experiment examining multiple paternity and mating dynamics in *Hapalochlaena maculosa*. The resulting genotype data generated were used for linkage map construction in this study. J.M.S.: Manuscript structure and aided with editing/proofing. I.R.C.: Manuscript structure and aided with editing/proofing. K.Z.: supervision, funding, results interpretation and editing/proofing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Australia and Pacific Science Foundation (APSF 14-9).

Institutional Review Board Statement: Ethical approval for all use and treatment of animals was approved by the James Cook University Animal Ethics Committee (Approval No. A1850). Collection of animals was conducted under Western Australia DPaW permit: SF010531 and Fisheries exemption: 26367.

Informed Consent Statement: Not applicable.

Data Availability Statement: DArTseq genotype data for all families included in this study have been made available in the Supplementary Materials.

Acknowledgments: We would like to thank Ross and Craig Cammilleri of the Fremantle Octopus Company for the use of their facilities, and both John dos Santos and Peter Stanitch for their help in obtaining animals from their by-catch. Additionally, thank you to Kirk Hardes and Timo Staeudchen for their concreting expertise and kayak skills.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sheumack, D.D.; Howden, M.E.H.; Spence, I. Occurrence of a Tetrodotoxin-like Compound in the Eggs of the Venomous Blue-Ringed Octopus (Hapalochlaena Maculosa). *Toxicon* **1984**, *22*, 811–812.
2. Mäthger, L.M.; Bell, G.R.R.; Kuzirian, A.M.; Allen, J.J.; Hanlon, R.T. How Does the Blue-Ringed Octopus (Hapalochlaena Lunulata) Flash Its Blue Rings? *J. Exp. Biol.* **2012**, *215*, 3752–3757.
3. White, J. Clinical Toxicology of Blue Ringed Octopus Bites. In *Handbook of: Clinical Toxicology of Animal Venoms and Poisons*; CRC Press: Boca Raton, FL, USA, 2018; pp. 171–175.
4. Jacups, S.P.; Currie, B.J. Blue-Ringed Octopuses: A Brief Review of Their Toxicology. *North. Territ. Nat.* **2008**, *20*, 50–57.
5. Whitelaw, B.L.; Strugnell, J.M.; Faou, P.; Da Fonseca, R.R.; Hall, N.E.; Norman, M.; Finn, J.; Cooke, I.R. Combined Transcriptomic and Proteomic Analysis of the Posterior Salivary Gland from the Southern Blue-Ringed Octopus and the Southern Sand Octopus. *J. Proteome Res.* **2016**, *15*, 3284–3297. [CrossRef]
6. Whitelaw, B.L.; Cooke, I.R.; Finn, J.; Da Fonseca, R.R.; Ritschard, E.A.; Gilbert, M.T.P.; Simakov, O.; Strugnell, J.M. Adaptive Venom Evolution and Toxicity in Octopods Is Driven by Extensive Novel Gene Formation, Expansion, and Loss. *Gigascience* **2020**, *9*, 1–15. [CrossRef]
7. Sohn, J.I.; Nam, J.W. The Present and Future of de Novo Whole-Genome Assembly. *Brief. Bioinform.* **2018**, *19*, 23–40. [CrossRef]
8. Albertin, C.B.; Simakov, O.; Mitros, T.; Wang, Z.Y.; Pungor, J.R.; Edsinger-Gonzales, E.; Brenner, S.; Ragsdale, C.W.; Rokhsar, D.S. The Octopus Genome and the Evolution of Cephalopod Neural and Morphological Novelties. *Nature* **2015**, *524*, 220–224. [CrossRef]
9. Kim, B.-M.; Kang, S.; Ahn, D.-H.; Jung, S.-H.; Rhee, H.; Yoo, J.S.; Lee, J.-E.; Lee, S.; Han, Y.-H.; Ryu, K.-B. The Genome of Common Long-Arm Octopus Octopus Minor. *Gigascience* **2018**, *7*, giy119.
10. Zarrella, I.; Herten, K.; Maes, G.E.; Tai, S.; Yang, M.; Seuntjens, E.; Ritschard, E.A.; Zach, M.; Styfhals, R.; Sanges, R. The Survey and Reference Assisted Assembly of the Octopus Vulgaris Genome. *Sci. Data* **2019**, *6*, 13.
11. Li, F.; Bian, L.; Ge, J.; Han, F.; Liu, Z.; Li, X.; Liu, Y.; Lin, Z.; Shi, H.; Liu, C.; et al. Chromosome-Level Genome Assembly of the East Asian Common Octopus (Octopus Sinensis) Using PacBio Sequencing and Hi-C Technology. *Mol. Ecol. Resour.* **2020**, *20*, 1572–1582. [CrossRef]
12. Gao, Y.; Natsukari, Y. Karyological Studies on Seven Cephalopods. *Jpn. J. Malacol.* **1990**, *49*, 126–145. [CrossRef]
13. Wang, J.; Zheng, X. Comparison of the Genetic Relationship between Nine Cephalopod Species Based on Cluster Analysis of Karyotype Evolutionary Distance. *Comp. Cytogenet.* **2017**, *11*, 477.
14. Leitwein, M.; Guinand, B.; Pouzadoux, J.; Desmarais, E.; Berrebi, P.; Gagnaire, P.A. A Dense Brown Trout (Salmo Trutta) Linkage Map Reveals Recent Chromosomal Rearrangements in the Salmo Genus and the Impact of Selection on Linked Neutral Diversity. *G3 Genes Genomes Genet.* **2017**, *7*, 1365–1376. [CrossRef]

15. Velmurugan, J.; Mollison, E.; Barth, S.; Marshall, D.; Milne, L.; Creevey, C.J.; Lynch, B.; Meally, H.; McCabe, M.; Milbourne, D. An Ultra-High Density Genetic Linkage Map of Perennial Ryegrass (*Lolium Perenne*) Using Genotyping by Sequencing (GBS) Based on a Reference Shotgun Genome Assembly. *Ann. Bot.* **2016**, *118*, 71–87. [CrossRef]
16. Choi, Y.; Kim, S.; Lee, J. Construction of an Onion (*Allium Cepa* L.) Genetic Linkage Map Using Genotyping-by-Sequencing Analysis with a Reference Gene Set and Identification of QTLs Controlling Anthocyanin Synthesis and Content. *Plants* **2020**, *9*, 616. [CrossRef]
17. Hagen, I.J.; Lien, S.; Billing, A.M.; Elgvin, T.O.; Trier, C.; Niskanen, A.K.; Tarka, M.; Slate, J.; Sætre, G.; Jensen, H. A Genome-wide Linkage Map for the House Sparrow (*Passer Domesticus*) Provides Insights into the Evolutionary History of the Avian Genome. *Mol. Ecol. Resour.* **2020**, *20*, 544–559. [CrossRef]
18. Manousaki, T.; Tsakogiannis, A.; Taggart, J.B.; Palaiokostas, C.; Tsaparis, D.; Lagnel, J.; Chatziplis, D.; Magoulas, A.; Papandroulakis, N.; Mylonas, C.C.; et al. Exploring a Nonmodel Teleost Genome through Rad Sequencing-Linkage Mapping in Common Pandora, Pagellus Erythrinus and Comparative Genomic Analysis. *G3 Genes Genomes Genet.* **2016**, *6*, 509–519. [CrossRef]
19. Duboule, D. The Rise and Fall of Hox Gene Clusters. *Development* **2007**, *134*, 2549–2560.
20. Lee, P.N.; Callaerts, P.; De Couet, H.G.; Martindale, M.Q. Cephalopod Hox Genes and the Origin of Morphological Novelties. *Nature* **2003**, *424*, 1061–1065. [CrossRef]
21. Callaerts, P.; Lee, P.N.; Hartmann, B.; Farfan, C.; Choy, D.W.Y.; Ikeo, K.; Fischbach, K.F.; Gehring, W.J.; Gert De Couet, H. HOX Genes in the Sepiolid Squid Euprymna Scolopes: Implications for the Evolution of Complex Body Plans. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 2088–2093. [CrossRef]
22. Biscotti, M.A.; Canapa, A.; Forconi, M.; Barucca, M. Hox and Parahox Genes: A Review on Molluscs. *Genesis* **2014**, *52*, 935–945.
23. Dawson, D.A.; Burke, T.; Hansson, B.; Pandhal, J.; Hale, M.C.; Hinten, G.N.; Slate, J. A Predicted Microsatellite Map of the Passerine Genome Based on Chicken-Passerine Sequence Similarity. *Mol. Ecol.* **2006**, *15*, 1299–1320. [CrossRef]
24. Voss, S.R.; Kump, D.K.; Walker, J.A.; Shaffer, H.B.; Voss, G.J. Thyroid Hormone Responsive QTL and the Evolution of Paedomorphic Salamanders. *Heredity* **2012**, *109*, 293–298. [CrossRef]
25. Sánchez-Molano, E.; Cerna, A.; Toro, M.A.; Bouza, C.; Hermida, M.; Pardo, B.G.; Cabaleiro, S.; Fernández, J.; Martínez, P. Detection of Growth-Related QTL in Turbot (Scophthalmus Maximus). *BMC Genomics* **2011**, *12*, 1–9. [CrossRef]
26. Estefanell, J.; Socorro, J.; Tuya, F.; Izquierdo, M.; Roo, J. Growth, Protein Retention and Biochemical Composition in Octopus Vulgaris Fed on Different Diets Based on Crustaceans and Aquaculture by-Products. *Aquaculture* **2011**, *322–323*, 91–98. [CrossRef]
27. Kawakami, T.; Smeds, L.; Backström, N.; Husby, A.; Qvarnström, A.; Mugal, C.F.; Olason, P.; Ellegren, H. A High-Density Linkage Map Enables a Second-Generation Collared Flycatcher Genome Assembly and Reveals the Patterns of Avian Recombination Rate Variation and Chromosomal Evolution. *Mol. Ecol.* **2014**, *23*, 4035–4058. [CrossRef]
28. Morse, P.; Huffard, C.L.; Meekan, M.G.; McCormick, M.I.; Zenger, K.R. Mating Behaviour and Postcopulatory Fertilization Patterns in the Southern Blue-Ringed Octopus, Hapalochlaena Maculosa. *Anim. Behav.* **2018**, *136*, 41–51. [CrossRef]
29. Sansaloni, C.; Petroli, C.; Jaccoud, D.; Carling, J.; Detering, F.; Grattapaglia, D.; Kilian, A. Diversity Arrays Technology (DArT) and next-Generation Sequencing Combined: Genome-Wide, High Throughput, Highly Informative Genotyping for Molecular Breeding of Eucalyptus. *BMC Proc.* **2011**, *5*, 1–2. [CrossRef]
30. Lind, C.E.; Kilian, A.; Benzie, J.A.H. Development of Diversity Arrays Technology Markers as a Tool for Rapid Genomic Assessment in Nile Tilapia, Oreochromis Niloticus. *Anim. Genet.* **2017**, *48*, 362–364. [CrossRef]
31. Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. CD-HIT: Accelerated for Clustering the next-Generation Sequencing Data. *Bioinformatics* **2012**, *28*, 3150–3152. [CrossRef]
32. Li, W.; Godzik, A. Cd-Hit: A Fast Program for Clustering and Comparing Large Sets of Protein or Nucleotide Sequences. *Bioinformatics* **2006**, *22*, 1658–1659. [CrossRef]
33. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; De Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575.
34. Quinlan, A.R.; Hall, I.M. BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features. *Bioinformatics* **2010**, *26*, 841–842.
35. Kalinowski, S.T.; Taper, M.L.; Marshall, T.C. Revising How the Computer Program CERVUS Accommodates Genotyping Error Increases Success in Paternity Assignment. *Mol. Ecol.* **2007**, *16*, 1099–1106. [CrossRef]
36. Jones, O.R.; Wang, J. COLONY: A Program for Parentage and Sibship Inference from Multilocus Genotype Data. *Mol. Ecol. Resour.* **2010**, *10*, 551–555. [CrossRef]
37. Rastas, P. Lep-MAP3: Robust Linkage Mapping Even for Low-Coverage Whole Genome Sequencing Data. *Bioinformatics* **2017**, *33*, 3726–3732. [CrossRef]
38. Li, H. Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997.
39. Catchen, J.; Amores, A.; Bassham, S. Chromonomer: A Tool Set for Repairing and Enhancing Assembled Genomes through Integration of Genetic Maps and Conserved Synteny. *G3 Genes Genomes Genet.* **2020**, *10*, 4115–4128. [CrossRef]
40. Finn, R.D.; Attwood, T.K.; Babbitt, P.C.; Bateman, A.; Bork, P.; Bridge, A.J.; Chang, H.-Y.; Dosztányi, Z.; El-Gebali, S.; Fraser, M.; et al. InterPro in 2017—beyond Protein Family and Domain Annotations. *Nucleic Acids Res.* **2016**, *45*, D190–D199. [CrossRef]
41. Slater, G.S.C.; Birney, E. Automated Generation of Heuristics for Biological Sequence Comparison. *BMC Bioinform.* **2005**, *6*, 1–11.

42. Rozewicki, J.; Li, S.; Amada, K.M.; Standley, D.M.; Katoh, K. MAFFT-DASH: Integrated Protein Sequence and Structural Alignment. *Nucleic Acids Res.* **2019**, *47*, W5–W10.
43. Stamatakis, A. RAxML-VI-HPC: Maximum Likelihood-Based Phylogenetic Analyses with Thousands of Taxa and Mixed Models. *Bioinformatics* **2006**, *22*, 2688–2690.
44. Yu, Y.; Zhang, X.; Yuan, J.; Li, F.; Chen, X.; Zhao, Y.; Huang, L.; Zheng, H.; Xiang, J. Genome Survey and High-Density Genetic Map Construction Provide Genomic and Genetic Resources for the Pacific White Shrimp Litopenaeus Vannamei. *Sci. Rep.* **2015**, *5*, 1–14. [CrossRef]
45. Chakravarti, A.; Lasher, L.K.; Reefer, J.E. A Maximum Likelihood Method for Estimating Genome Length Using Genetic Linkage Data. *Genetics* **1991**, *128*, 175–182. [CrossRef]
46. Postlethwait, J.H.; Johnson, S.L.; Midson, C.N.; Talbot, W.S.; Gates, M.; Ballinger, E.W.; Africa, D.; Andrews, R.; Carl, T.; Eisen, J.S.; et al. A Genetic Linkage Map for the Zebrafish. *Science* **1994**, *264*, 699–703. [CrossRef]
47. Danzmann, R.G. Linkage Analysis Package for Outcrossed Families with Male or Female Exchange of the Mapping Parent, Version 2.3. 2006. Available online: https://mybiosoftware.com/linkmfex-2-4-linkage-analysis-package-outcrossed-families-male-female-exchange-mapping-parent.html (accessed on 8 September 2017).
48. Hillis, D.M.; Bull, J.J. An Empirical Test of Bootstrapping as a Method for Assessing Confidence in Phylogenetic Analysis. *Syst. Biol.* **1993**, *42*, 182–192.
49. Semagn, K.; Bjørnstad, A.; Ndjiondjop, M.N. Principles, Requirements and Prospects of Genetic Mapping in Plants. *Afr. J. Biotechnol.* **2006**, *5*, 2569–2587.
50. Ferreira, A.; da Silva, M.F.; da Costa e Silva, L.; Cruz, C.D. Estimating the Effects of Population Size and Type on the Accuracy of Genetic Maps. *Genet. Mol. Biol.* **2006**, *29*, 187–192. [CrossRef]
51. Jones, D.B.; Jerry, D.R.; Khatkar, M.S.; Raadsma, H.W.; Zenger, K.R. A High-Density SNP Genetic Linkage Map for the Silver-Lipped Pearl Oyster, Pinctada Maxima: A Valuable Resource for Gene Localisation and Marker-Assisted Selection. *BMC Genomics* **2013**, *14*, 810. [CrossRef]
52. Adema, C.M.; Hillier, L.W.; Jones, C.S.; Loker, E.S.; Knight, M.; Minx, P.; Oliveira, G.; Raghavan, N.; Shedlock, A.; Do Amaral, L.R. Whole Genome Analysis of a Schistosomiasis-Transmitting Freshwater Snail. *Nat. Commun.* **2017**, *8*, 15451.
53. Takeuchi, T. Molluscan Genomics: Implications for Biology and Aquaculture. *Curr. Mol. Biol. Rep.* **2017**, *3*, 297–305. [CrossRef]
54. Stapley, J.; Feulner, P.G.D.; Johnston, S.E.; Santure, A.W.; Smadja, C.M. Variation in Recombination Frequency and Distribution across Eukaryotes: Patterns and Processes. *Philos. Trans. R. Soc. B Biol. Sci.* **2017**, *372*, 20160455. [CrossRef]
55. Tiley, G.P.; Burleigh, G. The Relationship of Recombination Rate, Genome Structure, and Patterns of Molecular Evolution across Angiosperms. *BMC Evol. Biol.* **2015**, *15*, 1–14. [CrossRef]
56. Boissinot, S.; Entezam, A.; Furano, A.V. Selection against Deleterious LINE-1-Containing Loci in the Human Lineage. *Mol. Biol. Evol.* **2001**, *18*, 926–935. [CrossRef]
57. Bartolomé, C.; Maside, X.; Charlesworth, B. On the Abundance and Distribution of Transposable Elements in the Genome of Drosophila Melanogaster. *Mol. Biol. Evol.* **2002**, *19*, 926–937. [CrossRef]
58. Fu, H.; Zheng, Z.; Dooner, H.K. Recombination Rates between Adjacent Genic and Retrotransposon Regions in Maize Vary by 2 Orders of Magnitude. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 1082–1087. [CrossRef]
59. Nakamura, H. A Review of Molluscan Cytogenetic Information Based on the CISMOCH: Computerized Index System for Molluscan Chromosomes: Bivalvia, Polyplacophora and Cephalopoda. *Venus Jpn. J. Malacol.* **1985**, *44*, 193–225. [CrossRef]
60. Amarasinghe, S.L.; Su, S.; Dong, X.; Zappia, L.; Ritchie, M.E.; Gouil, Q. Opportunities and Challenges in Long-Read Sequencing Data Analysis. *Genome Biol.* **2020**, *21*, 1–16.
61. Simakov, O.; Marlétaz, F.; Yue, J.X.; O'Connell, B.; Jenkins, J.; Brandt, A.; Calef, R.; Tung, C.H.; Huang, T.K.; Schmutz, J.; et al. Deeply Conserved Synteny Resolves Early Events in Vertebrate Evolution. *Nat. Ecol. Evol.* **2020**, *4*, 820–830. [CrossRef]
62. Fierst, J.L. Using Linkage Maps to Correct and Scaffold de Novo Genome Assemblies: Methods, Challenges, and Computational Tools. *Front. Genet.* **2015**, *6*, 220.
63. Hedgecock, D.; Shin, G.; Gracey, A.Y.; van den Berg, D.; Samanta, M.P. Second-Generation Linkage Maps for the Pacific Oyster Crassostrea Gigas Reveal Errors in Assembly of Genome Scaffolds. *G3 Genes Genomes Genet.* **2015**, *5*, 2007–2019. [CrossRef]
64. Shaw, K.L.; Lesnick, S.C. Genomic Linkage of Male Song and Female Acoustic Preference QTL Underlying a Rapid Species Radiation. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 9737–9742. [CrossRef]
65. Han, S.; Yuan, M.; Clevenger, J.P.; Li, C.; Hagan, A.; Zhang, X.; Chen, C.; He, G. A Snp-Based Linkage Map Revealed QTLs for Resistance to Early and Late Leaf Spot Diseases in Peanut (*Arachis Hypogaea* L.). *Front. Plant Sci.* **2018**, *9*, 1012. [CrossRef]
66. Zhang, Z.S.; Hu, M.C.; Zhang, J.; Liu, D.J.; Zheng, J.; Zhang, K.; Wang, W.; Wan, Q. Construction of a Comprehensive PCR-Based Marker Linkage Map and QTL Mapping for Fiber Quality Traits in Upland Cotton (*Gossypium Hirsutum* L.). *Mol. Breed.* **2009**, *24*, 49–61. [CrossRef]
67. Wang, L.; Chua, E.; Sun, F.; Wan, Z.Y.; Ye, B.; Pang, H.; Wen, Y.; Yue, G.H. Mapping and Validating QTL for Fatty Acid Compositions and Growth Traits in Asian Seabass. *Mar. Biotechnol.* **2019**, *21*, 643–654. [CrossRef]
68. Berger, E. *Aquaculture of Octopus Species: Present Status, Problems and Perspectives*; University of Plymouth: Plymouth, UK, 2010.
69. Vaz-Pires, P.; Seixas, P.; Barbosa, A. Aquaculture Potential of the Common Octopus (Octopus Vulgaris Cuvier, 1797): A Review. *Aquaculture* **2004**, *238*, 221–238. [CrossRef]

70. Shetty, S.; Griffin, D.K.; Graves, J.A.M. Comparative Painting Reveals Strong Chromosome Homology over 80 Million Years of Bird Evolution. *Chromosom. Res.* **1999**, *7*, 289–295. [CrossRef]
71. Van Tuinen, M.; Hedges, S.B. Calibration of Avian Molecular Clocks. *Mol. Biol. Evol.* **2001**, *18*, 206–213. [CrossRef]
72. Strasburg, J.L.; Sherman, N.A.; Wright, K.M.; Moyle, L.C.; Willis, J.H.; Rieseberg, L.H. What Can Patterns of Differentiation across Plant Genomes Tell Us about Adaptation and Speciation? *Philos. Trans. R. Soc. B Biol. Sci.* **2012**, *367*, 364–373.
73. Price, N.; Moyers, B.T.; Lopez, L.; Lasky, J.R.; Grey Monroe, J.; Mullen, J.L.; Oakley, C.G.; Lin, J.; Ågren, J.; Schrider, D.R.; et al. Combining Population Genomics and Fitness QTLs to Identify the Genetics of Local Adaptation in Arabidopsis Thaliana. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 5028–5033. [CrossRef]