

SCIENTIFIC REPORTS



OPEN

Construction of a high-density genetic map and QTL mapping for pearl quality-related traits in *Hyriopsis cumingii*

Zhi-Yi Bai^{1,2,*}, Xue-Kai Han^{1,*}, Xiao-Jun Liu¹, Qing-Qing Li¹ & Jia-Le Li^{1,2,3}

Received: 05 February 2016
Accepted: 11 August 2016
Published: 02 September 2016

A high-density genetic map is essential for quantitative trait locus (QTL) fine mapping. In this study, 4,508 effective single nucleotide polymorphism markers (detected using specific-locus amplified fragment sequencing) and 475 microsatellites were mapped to 19 linkage groups (LGs) using a family with 157 individuals. The map spanned 2,713 cM, with an average of 259 markers and 79 loci per LG and an average inter-marker distance of 1.81 cM. To identify QTLs for pearl quality traits, 26 putatively significant QTLs were detected for 10 traits, including, three for shell width, seven for body weight, two for shell weight, two for margin mantle weight, five for inner mantle weight, and seven for shell nacre colour. Among them, five QTLs associated with shell nacre colour were mapped to LG17 and explained 19.7% to 22.8% of the trait variation; this suggests that some important genes or loci determine shell nacre colour in LG17. The linkage map and mapped QTLs for shell nacre colour would be useful for improving the quality of *Hyriopsis cumingii* via marker-assisted selection.

The triangle sail mussel, *Hyriopsis cumingii*, is widely distributed in large lakes, rivers, and estuaries in China. Nearly 95% of freshwater pearls worldwide are produced in China, of which about 70% are produced by *H. cumingii*¹. Currently, modern *H. cumingii* aquaculture has several issues such as degradation of germplasm resources and decline in pearl quality. Therefore, it is important to genetically improve *H. cumingii*. Although traditional breeding has played a crucial role in genetic improvement, it relies mainly on phenotype, which is easily affected by environmental factors². Nowadays, marker-assisted selection (MAS) provides a very efficient tool for selective breeding based on genotype. Unlike traditional selective breeding based on phenotype, MAS offers precise and rapid selection of the targeted gene and accelerates breeding. Development of effective genes or loci that affect traits is a precondition for MAS. Karyotype analysis has shown that *H. cumingii* has 38 chromosomes (19 pairs)³, and its genome size has been estimated to be approximately 3 Gb on the basis of k-mer analysis. The larger genome size, higher heterozygosity, and redundant sequences resulted in a lag in developing gene or marker resources for *H. cumingii*. Rapid development of sequencing techniques provided us with a great opportunity to produce valuable genomic resources for *H. cumingii*.

Quantitative trait locus (QTL) mapping is an effective method for identifying markers associated with economic traits for use in MAS, and construction of genetic linkage maps is a decisive precondition^{4,5}. To date, linkage maps have been constructed for more than 40 aquaculture species, and QTL mapping has been performed for over 20 aquaculture species⁶. Nevertheless, QTL mapping has not been widely used for genetic breeding of aquaculture species. A key reason is that most first-generation genetic maps for aquaculture species have a limited number of amplified fragment length polymorphisms, restricted fragment length polymorphisms, or simple sequence repeats (SSRs); these maps have a low density, making it difficult to perform fine-scale mapping to narrow down genomic regions strongly associated with important traits. Therefore, a sufficient number of high-quality molecular markers are crucial for constructing a high-density linkage map and determining the efficiency and veracity of the genetic map⁷.

¹Key Laboratory of Freshwater Aquatic Genetic Resources, Shanghai Ocean University, Ministry of Agriculture, Shanghai 201306, China. ²Shanghai Engineering Research Centre of Aquaculture, Shanghai Ocean University, Shanghai 201306, China. ³Aquaculture Division, E-Institute of Shanghai Universities, Shanghai Ocean University, Shanghai 201306, China. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to J.-L.L. (email: jlli2009@126.com)

Type	No. of reads	No. of SLAFs	Coverage	Ratio
Polymorphic SLAF	39,152,380	132,542	295	55.3%
Non-polymorphic SLAF	29,278,441	107,162	273	44.7%
Total	68,430,821	239,704	285	100.0%

Table 1. Mining results for SLAF markers.

Single nucleotide polymorphism (SNP) is a variation in just one nucleotide at the genome level⁸. When compared with previous molecular markers, SNPs possess advantages such as high abundance in the genome and heritable stability. In particular, SNPs are suitable for automated large-scale genotyping, allowing for genotyping of a large number of cost-effective markers to construct high-density genetic maps⁹. With the rapid development of sequencing techniques in the past decade, especially next-generation sequencing (NGS) techniques, various genotyping methods have been developed and thousands of SNPs were identified in many species. Currently, more than 30 different methods are being used for SNP genotyping¹⁰. For instance, restriction site-associated sequencing (RAD-seq) is being widely used because of its rapidity and reliability in genotyping and polymorphism identification¹¹, and several methods with simpler library preparation protocols, such as 2b-RAD¹² and double digest RAD-seq¹³, have been developed. Elshire *et al.*¹⁴ introduced an efficient genotyping-by-sequencing approach that can simultaneously detect and score thousands of SNPs. Besides genotyping based on NGS, SNPs are usually genotyped with SNP arrays. Some SNPs arrays have been developed and used for aquaculture species, such as the Atlantic salmon¹⁵, catfish¹⁶, common carp¹⁷, Atlantic cod¹⁸, and rainbow trout¹⁹. Because of these effective approaches, high-density SNP-based linkage maps have been constructed for several aquaculture species, such as the Atlantic salmon²⁰, rainbow trout²¹, channel catfish²², Japanese flounder²³, and Zhikong scallop⁷. Recently, a more effective method called specific-locus amplified fragment sequencing (SLAF-seq) was proposed by Sun *et al.*²⁴. SLAF-seq is used a pre-designed reduced representation strategy and deep sequencing to improve the efficiency and accuracy of genotyping. Moreover, it provides a double barcode system that can be applied to large populations and is needed to construct the SLAF-seq library. After high-throughput sequencing, a large number of sequences were collected and used to obtain polymorphism SLAF tags by software alignment; then, specific SNP sites were found on the basis of these polymorphism SLAF tags. Thus, SLAF and SNP markers obtained via this approach are essentially the same. This technology has been applied for the construction of high-density linkage maps for several species, such as cucumber²⁵, soybean²⁶, common carp²⁴, and tea²⁷.

Since more genetic maps have been constructed for bivalves, QTL mapping for some important economic traits has been performed recently. For instance, QTL mapping for growth traits and disease resistance has been conducted in the Pacific oyster^{28,29}; QTL mapping for growth and shell colour has been conducted in *Argopecten irradians*^{30,31}; and QTL mapping for growth traits and sex-determination has been conducted in *Chlamys farreri*⁷. In addition, for some pearl mussels, QTLs have been identified for growth and shell colour in the freshwater mussel *H. cumingii*³², and QTLs for pearl quality and shell growth have been identified in *Pinctada maxima*^{33,34}. The identified QTLs will provide the basis for MAS of important economic traits in these bivalves.

In this study, we used 157 F₁ individuals to perform whole-genome genotyping with SLAF-seq and constructed a high-density genetic map for *H. cumingii*. Furthermore, we performed QTL screening for pearl quality-related traits by using this map. The high-density genetic map will provide an essential tool for QTL fine mapping and identifying candidate genes associated with complex traits. Identification of QTLs related to pearl quality traits may remarkably facilitate breeding programs in the future.

Results

Analysis of SLAF-seq data for 157 F₁ individuals and two parents. In this study, we first constructed the SLAF-seq library for *H. cumingii* with 157 F₁ individuals and two parents. Using high-throughput sequencing and bioinformatics, a total of 139,113,148 pair-end reads were obtained for the SLAF library. Among them, 88.64% bases were of a high quality (Q score > 30), and GC content was 37.45% on average. The number of reads for male and female parents was 9,025,115 and 8,571,660, respectively. For the F₁ population, read numbers ranged from 523,546 to 1,009,487 (average, 773,990; Table S1).

All reads were clustered into SLAFs on the basis of sequence similarity. After correcting or discarding low-depth SLAF tags, 239,704 high-quality SLAFs were identified, of which 201,805 and 206,806 were detected in the male and female parents, respectively. The coverage of SLAFs for paternal and maternal parents was 19.25-fold and 20.70-fold, respectively. For the F₁ population, 62,790 to 140,362 SLAFs (average, 123,038) were generated, indicating that the coverage of SLAFs in each progeny ranged from 2.31-fold to 3.61-fold (average, 3.11-fold; Fig. S1).

SLAF marker detection and genotype definition. Among the 239,704 high-quality SLAFs, 132,542 were polymorphic with a polymorphism percentage of 55.3% (Table 1). Of the 132,542 polymorphic SLAFs, 108,004 were classified into eight segregation patterns. The segregation patterns were as follows: ab × cd, ef × eg, hk × hk, lm × ll, nn × np, aa × bb, ab × cc, and cc × ab (Fig. 1). Since the F₁ population was used in this study, only SLAF markers with segregation patterns of ab × cd, ef × eg, hk × hk, lm × ll, nn × np, ab × cc, or cc × ab were used for the genetic map construction. To further improve SLAF accuracy, only SLAFs with more than 80% integrity and more than 20-fold average sequence depths in the parents were used for the map construction. Finally, 4,508 of the 108,004 markers were selected for the linkage map construction (Table 2). Of these 4,508 markers, 60.96% were heterozygous in the male parent; 61.78%, in the female parent; and 21.33% to 47.07%, in the F₁ individuals.

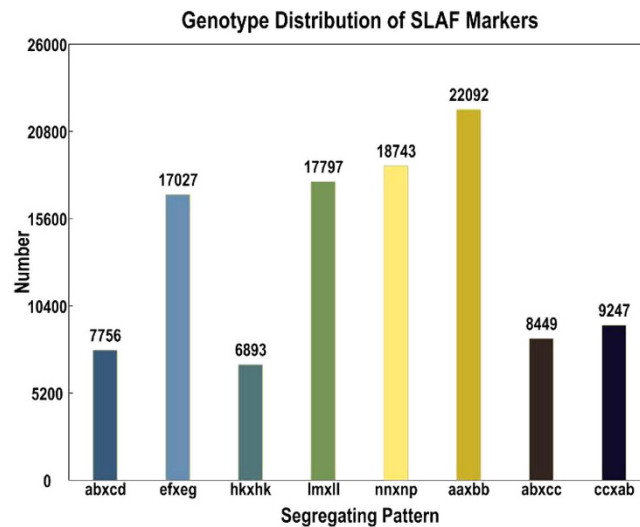


Figure 1. Distribution of SLAF markers in eight segregation patterns.

Type	No. of SLAFs	Percentage
abxcd	52	1.15%
efzeg	387	8.58%
hkxhk	586	13.00%
lmxll	1,723	38.22%
nnxnp	1,760	39.04%
Total	4,508	100.00%

Table 2. Statistics for segregation types of the SLAF markers.

Average sequencing depths of the 4,508 markers were 86.71-fold and 92.66-fold for the male and female parents, respectively, and 7.99-fold for each progeny.

Construction of the high-density linkage map. To construct a high-density, high-quality genetic linkage map for *H. cumingii*, the newly developed 4,508 SNPs and 506 SSR markers from the framework map³² were used. After linkage analysis, 4,920 markers were mapped to 19 linkage groups (Table 3, Fig. 2). The final map was 2,713 cM in length, with an average of 79 loci per LG and an average inter-marker distance of 1.81 cM. The number of markers mapped to each linkage group varied from 114 markers in LG15 to 436 markers in LG2, with an average of 258 markers per LG. The largest LG was LG2, which contained 436 markers with a genetic length of 196.50 cM, and the smallest LG was LG10, which contained only 180 markers with a length of 91.09 cM. The mean chromosome region length was 142.80 cM. The ‘Gap ≤ 5 cM’ value (indicates the percentages of gaps in which the distance between adjacent markers is smaller than 5 cM) of the 19 linkage groups ranged from 80.28% to 100% (average, 97.79%). The maximum gap was 18.91 cM, which was observed in LG15.

In this study, sex-specific maps were also constructed (Table S4). The male map contained 3,233 markers and spanned 2,561 cM, while the female map contained 3,123 markers and spanned 2,810 cM. In addition, the male map contained 170 markers and 48 loci per LG, while the female map contained 164 markers and 56 loci per LG. The average inter-marker distances for the male and female maps were 2.78 and 2.66 cM, respectively. Although the lengths of the male and female maps were similar, we observed significant differences in some of LGs. For instance, the ratios of lengths of LG10, LG13, and LG18 in the male and female maps were 1.39, 1.45, and 1.61, respectively (Table S4).

Comparison of the constructed map with the SSR-based framework map. Since we constructed an SSR-based framework map for *H. cumingii* in a previous study, we performed a brief comparison of the two maps. Although both maps comprised 19 linkage groups, significant differences were observed in the number of markers, map length, and resolution between the two maps. The SSR-based framework map contained only 506 SSR markers and spanned 1,922.3 cM, which is obviously less than the 4,920 markers and 2,713 cM map length observed in this study. Our current linkage map has higher resolution at 1.81 cM, which is significantly superior to the resolution of the previous genetic map with an average inter-marker distance of 3.99 cM. Besides, we noted many large gaps in the SSR-based map and a lesser number of large gaps in the current map.

QTL mapping for pearl quality-related traits. The basic statistics for the pearl quality traits are listed in Table S2. On the basis of the high-density genetic map, a total of 26 QTLs were detected on LG1, LG4, LG8,

Linkage group	Marker no.			Number of loci	Size (cM)	Average distance (cM)	Gap \leq 5	Max gap (cM)
	SSR	SNP	Total					
LG01	23	303	326	84	152.06	1.81	92.77	15.99
LG02	36	400	436	127	196.50	1.55	95.23	8.09
LG03	35	387	422	105	157.35	1.50	97.12	13.53
LG04	36	269	305	72	185.50	2.58	80.28	17.24
LG05	29	202	231	81	152.51	1.88	95.00	15.34
LG06	33	219	252	82	155.03	1.89	91.36	14.60
LG07	27	184	211	64	102.17	1.60	95.24	14.71
LG08	24	227	251	86	168.70	1.96	92.94	16.44
LG09	24	159	183	65	124.08	2.00	90.63	13.26
LG10	35	145	180	47	91.09	1.94	93.48	16.21
LG11	27	340	367	98	176.70	1.80	92.78	17.56
LG12	21	337	358	85	102.09	1.20	100.00	4.91
LG13	11	243	254	89	121.64	1.37	96.60	8.73
LG14	24	230	254	89	151.57	1.70	95.45	10.55
LG15	23	91	114	47	139.16	2.96	89.13	18.91
LG16	23	212	235	82	165.53	2.04	90.12	12.61
LG17	15	154	169	62	99.13	1.60	98.36	7.55
LG18	14	173	187	54	137.82	2.55	90.91	11.00
LG19	15	170	185	81	134.55	1.66	93.75	10.07
Total	475	4,445	4,920	1,500	2,713.17	/	/	/
Average	25	234	259	79	142.80	1.81	93.46	/

Table 3. Summary of the integrated linkage map for *H. cumingii*.

LG14, LG15, and LG17 (Table 4, Fig. 3). Phenotypic variability explained by each QTL ranged from 8.4% to 22.8% (average, 12.3%). Three QTLs for shell width were located on LG1, and the confidence intervals were at 0–11.3, 15.5, and 21.8–22.4 cM. Two QTLs for shell weight were detected at 15.5 cM on LG1 and 11.2 cM on LG15. Seven QTLs for body weight spanned discontinuously from 76.3 to 157.6 cM on LG8. For inner mantle weight, three QTLs spanned discontinuously from 107.6 to 184.5 cM on LG4, and two QTLs were at 15.5 and 21.8–27.6 cM on LG1. Besides, two QTLs for margin mantle weight were located on LG8, and the confidence intervals were at 78.6–79.3 and 82.9 cM. For the growth-related traits, three QTLs (Wi-2, Whi-4, and Sw-1) were found in the same region on LG1, and they showed a percentage of phenotypic variance explained (PVE) of 10.0%, 8.4%, and 8.8%, respectively. Two QTLs (We-2 and Who-2) were found in the same region on LG8 near Marker93674 and Marker261215, with a PVE of 9.7% and 11.8%, respectively. Additionally, seven QTLs associated with shell nacre colour were detected, showing a high percentage of PVE of 19.7% to 22.8%. Of these seven QTLs, five significant QTLs related to shell nacre colour tended to colocalize, clustering in a special region (17.3 cM to 36.1 cM) on LG17 (Fig. 3).

The validation of four QTLs associated with shell nacre colour on LG17. Statistical analysis showed that all three markers were significantly associated with shell nacre colour ($P < 0.05$). For Marker257873, the *H. cumingii* with CC genotype showed a significantly higher value in AL while significantly lower values in Aa, Ab and AdE, compared with the individuals with CT or TT genotypes. For Marker189034, the individuals with AA genotype showed significantly higher values in Aa, Ab and AdE, but a significantly lower value in AL when compared with the individuals with GA or GG genotypes. Similarly, as for Marker273509, the individuals with GG genotype had a significantly lower value in AL but higher values in Aa, Ab and AdE (Table 5).

Discussion

High-density linkage maps are especially important for QTL mapping, positional cloning, comparative genomics, MAS, and identification of functional genes^{35–38}. Currently, although SNPs have become the first-selected genetic markers for high-density linkage map construction, their widespread application remains limited by the high costs of high-throughput sequencing and genotyping³⁹. Fortunately, NGS-based marker identification and genotyping technologies provide a powerful method for the identification of large numbers of SNPs. The SLAF-seq approach balances genotyping accuracy and sequencing cost, thus providing an economical and efficient method for linkage mapping of non-model species with complex genomes^{24–27,40}.

In this study, SLAF-seq was used to develop SLAF markers in an F_1 full-sib family of *H. cumingii*. By SLAF library construction and high-throughput sequencing, 239,704 high-quality SLAFs were detected, and 129,967 SLAFs were polymorphic with a high polymorphism rate of 54.22%. Although the SLAF-seq strategy provides a powerful tool for large-scale SNP identification, some inevitable errors and missing values occur during SLAF marker development and genotyping⁴¹. Therefore, strict criteria were used to avoid false-positive markers, and 4,508 markers were finally used for the linkage mapping. Since the sequencing depth reflects the accuracy of *de novo* SNP identification and genotyping quality scores can be used to detect suspicious markers, they are

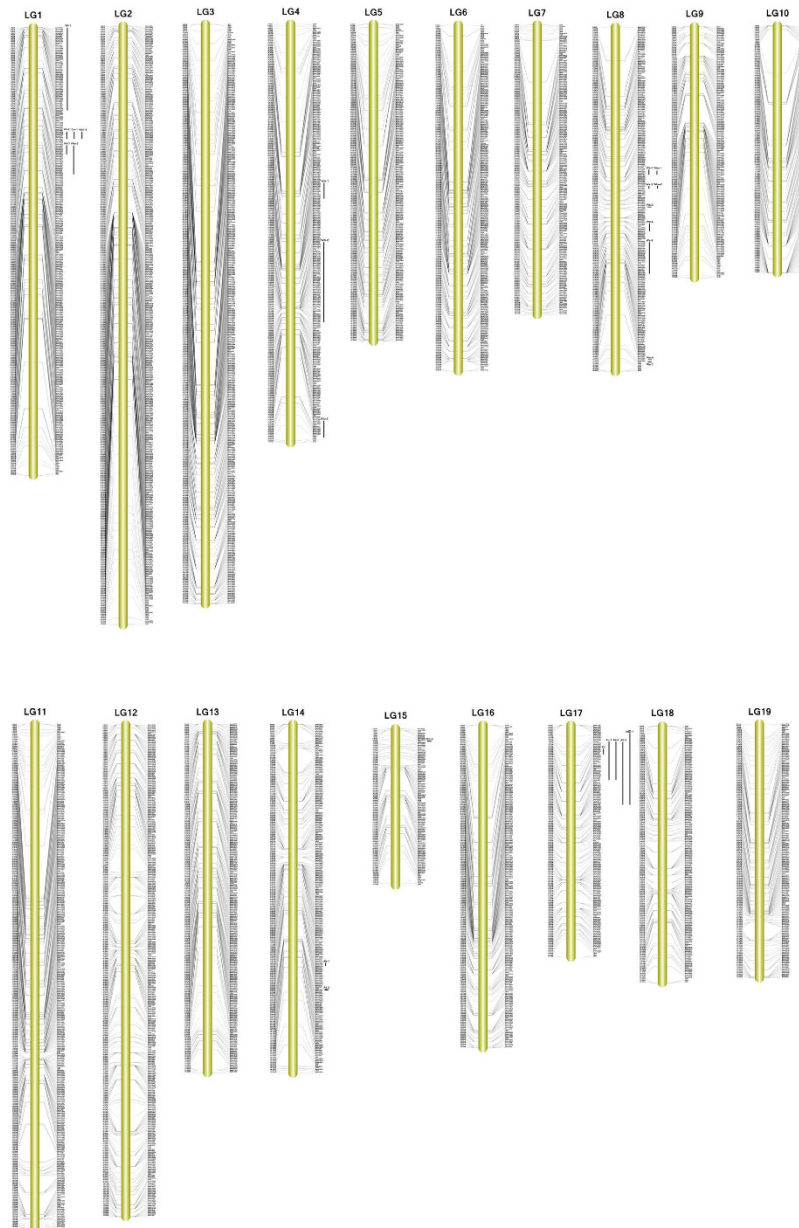


Figure 2. The high-density genetic map and QTLs for traits associated with pearl quality in *Hyriopsis cumingii*. The markers and their locations are shown on the right and left sides, respectively. QTLs are marked on the right of each linkage map.

both important criteria for ensuring genotyping accuracy. The minimum sequencing depth for each individual is 6-fold when the SLAF-seq approach is used²⁴. In our study, average sequencing depths were 86.71-fold and 92.66-fold for the male and female parents and nearly 8-fold for each progeny, which provided a high level of genotyping accuracy. In addition, genotyping quality scores of these 4,508 markers reached 30, which effectively eliminated the suspicious markers. We excluded markers with significant segregation distortion (the genotype integrity of these markers was less than 80% in the mapping population), thus ensuring the quality of the SLAF markers and genotyping accuracy. Therefore, our SNP genotyping data can be reliably used for high-density linkage map construction.

We constructed a high-density genetic map for *H. cumingii*, containing 4,920 markers with a resolution of 1.81 cM. The high-density genetic map is a great improvement over our previous map for *H. cumingii*³², providing an important tool for fine mapping of QTLs for important traits. Our current linkage map is better than previous genetic maps for bivalves such as *P. maxima* (1,189 SNP markers)⁴², *Pinctada fucata martensii* (3,117 markers)⁴³, and *C. farreri* (3,806 markers)⁷ with respect to marker number, but it is slightly inferior with respect to average marker interval because of the larger genome size of *H. cumingii*. We also observed some large gaps in the current map; this suggested that there may be a higher level of recombination or that markers in these regions have not

Trait	QTL ID	LOD threshold	LG group	QTL interval (cM)	Peak position (cM)	Peak LOD value	PVE (%)		Marker number
							Minimum	Maximum	
Shell width	Wi-1	3.0	LG1	0–11.3	1.6	3.84	9.1	10.7	62
	Wi-2	3.0	LG1	15.5	15.5	3.58	10.0	10.0	6
	Wi-3	3.0	LG1	21.8–22.4	21.8	3.14	8.5	8.8	7
Body weight	We-1	3.0	LG8	76.3–79.6	79.6	3.25	8.4	9.1	4
	We-2	3.0	LG8	82.9	82.9	3.47	9.7	9.7	2
	We-3	3.0	LG8	89.2	89.2	3.06	8.8	8.8	1
	We-4	3.0	LG8	95.1	95.1	3.31	9.2	9.2	6
	We-5	3.0	LG8	99.6–105.7	100.9	3.60	9.3	10.0	23
	We-6	3.0	LG8	155.6	155.6	3.35	9.5	9.5	1
	We-7	3.0	LG8	157.6	157.6	3.04	8.5	8.5	1
Shell weight	Sw-1	3.0	LG1	15.5	15.5	3.12	8.7	8.8	6
	Sw-2	3.0	LG15	11.2	11.2	3.12	8.7	8.7	1
Inner mantle weight	Whi-1	3.0	LG4	107.6–108.2	108.2	3.04	8.4	8.5	12
	Whi-2	3.0	LG4	120.9–126.1	125.8	3.27	8.5	9.8	59
	Whi-3	3.0	LG4	148.0–184.5	167.8	3.76	8.6	10.4	13
	Whi-4	3.0	LG1	15.5	15.5	3.01	8.4	8.4	6
	Whi-5	3.0	LG1	21.8–27.6	27.6	3.22	8.4	9.2	22
Margin mantle weight	Who-1	4.1	LG8	78.6–79.3	78.6	4.25	11.4	11.7	2
	Who-2	4.1	LG8	82.9	82.9	4.28	11.8	11.8	2
Purple mantle scar length	P-1	7.2	LG17	24.7	24.7	8.06	21.0	21.1	5
AL	AL-1	6.0	LG17	17.3–36.1	28.5	7.39	16.5	19.7	29
Aa	Aa-1	6.1	LG17	17.3–36.1	28.5	8.77	16.4	22.8	29
Ab	Ab-1	6.6	LG14	79.6	79.6	6.72	17.9	17.9	2
	Ab-2	6.6	LG14	85.0	85.0	6.72	17.9	17.9	1
	Ab-3	6.6	LG17	19.4–40.1	28.5	8.55	18.6	22.2	45
AdE	AdE-1	5.9	LG17	4.7–40.1	24.8	8.79	15.9	22.8	52

Table 4. Summary of QTLs for pearl quality-related traits in F₁ populations of *H. cumingii*.

been developed. Hence, other enzymes may be used for SLAF library preparation, and the size of the mapping population needs to be improved in the future.

When compared with the SSR-based framework map³², the map constructed in this study spanned a cumulative distance of 2,713 cM, showing a significant increase in map length. Of course, the genetic length of a linkage map could be affected by many factors, such as the number and type of markers, size of the mapping population, and accuracy of the genotyping³⁷. Errors in genotyping usually cause elongation of a linkage map. In our study, strict criteria were used during SLAF marker development and genotyping, which ensured the quality of the SLAF markers and greatly reduced genotyping errors; this enhanced the accuracy of the genotyping and, thus, effectively avoided artificial inflation of map length. Therefore, elongation of the genetic map in this study was probably caused by the variation in the number and type of the markers and changes in the size of the mapping population.

Accuracy of marker order and the genetic map is one of the most interesting aspects for people who work with genetic mapping. In this study, the order of common SSR markers in the present map and the previous framework map was compared³². For most LGs, a roughly similar order was observed for the SSR markers; however, different orders of common SSR markers were observed in some LGs. The variation may be partly due to the different numbers of markers and individuals used for the two genetic maps. A previous study showed that a larger-sized population is considered sufficient for the construction of reasonably accurate genetic maps⁴⁴. Furthermore, this effect will be even more obvious when the number of different families increased, since it could increase the power to map loci by increasing the number of informative meiosis events. In addition, different methods for map construction may influence the accuracy of marker order and the genetic map. Some minor differences in marker order have been observed between linkage maps constructed using JoinMap4.1 and HighMap⁴¹. A previous study has reported that HighMap has many advantages over JoinMap4.1⁴⁵. HighMap not only permits the use of more NGS data but could also guarantee higher accuracy of marker order and map distance for data with a large proportion of missing or erroneous markers.

Many factors affect pearl quality, such as size, shape, colour, and lustre. Since production of a non-nucleated pearl requires the participation of both donor and host mussels, previous studies have highlighted the link between pearl traits and growth-related traits in *H. cumingii*. A previous study has shown that shell weight, body weight, and shell width are significantly associated with pearl diameter and pearl yield⁴⁶. Margin mantle of the donor mussel is used to make the saibo, and size and thickness of the margin mantle determines the size and thickness of the saibo. The inner mantle, in which a non-nucleated pearl is implanted in the receptor mussel, directly affects pearl growth. Therefore, these five growth traits are considered related to pearl quality⁴⁷. Pearl

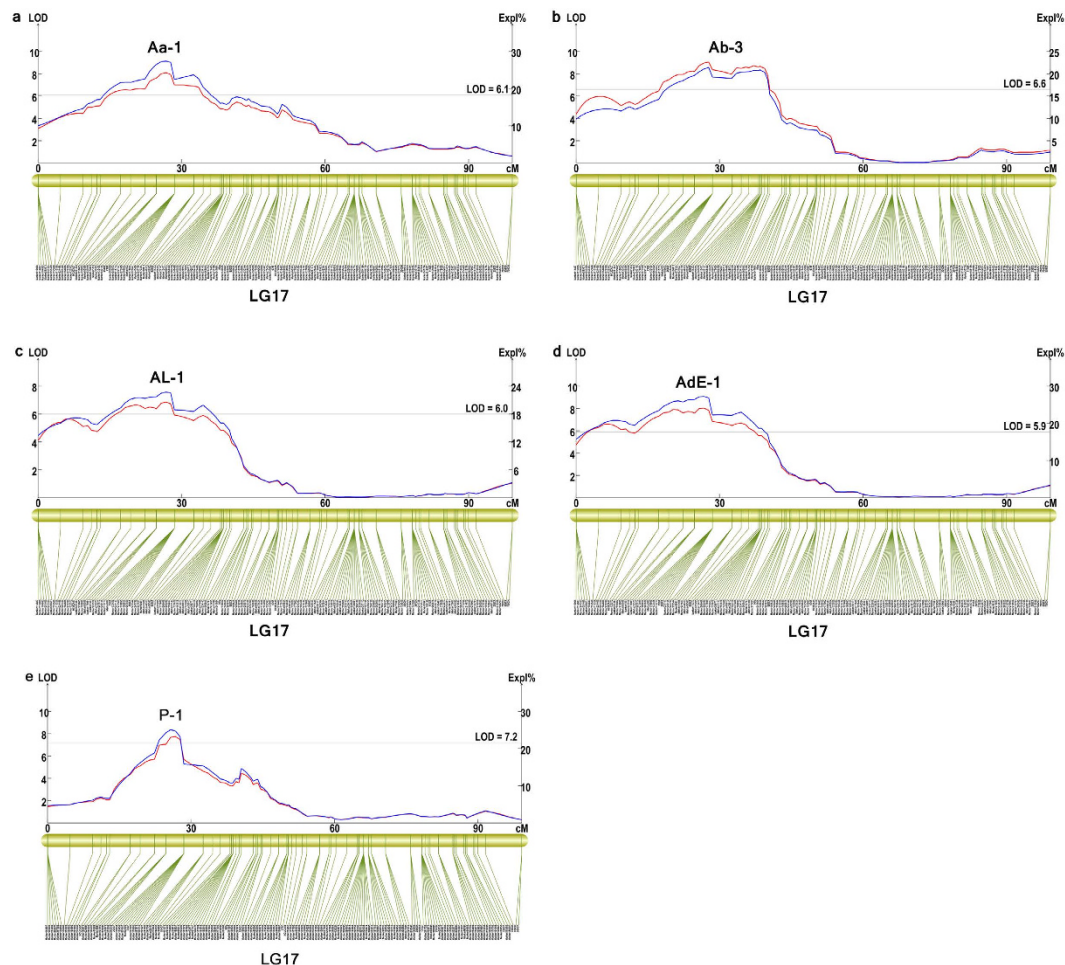


Figure 3. Mapping of significant QTLs for shell nacre colour. (a) QTL for Aa in LG17; (b), QTL for Ab in LG17; (c), QTL for AL in LG17; (d), QTL for AdE in LG17; (e), QTL for purple mantle scar length in LG17. The blue line and red line represent the LOD score and the explained phenotypic variation of QTL, respectively.

SNP	Genotype	No.	AL	Aa	Ab	AdE
Marker189034	AA	86	49.81 ± 5.36 ^a	4.48 ± 2.47 ^a	-6.30 ± 4.41 ^a	51.96 ± 5.18 ^a
	GA	144	52.38 ± 6.33 ^b	3.48 ± 2.98 ^b	-7.70 ± 4.18 ^b	49.23 ± 6.33 ^b
	GG	48	52.64 ± 5.38 ^b	3.81 ± 2.39 ^{Ab}	-8.64 ± 4.48 ^b	49.12 ± 5.24 ^b
<i>P</i> -value			0.0029 [*]	0.0274 [*]	0.0064 [*]	0.0015 [*]
Marker257873	CC	85	53.98 ± 5.71 ^a	2.92 ± 2.40 ^a	-8.60 ± 4.72 ^a	47.74 ± 5.60 ^a
	CT	119	50.89 ± 6.19 ^b	4.21 ± 2.70 ^b	-7.17 ± 4.01 ^b	50.75 ± 6.07 ^b
	TT	68	50.33 ± 5.36 ^b	4.02 ± 2.81 ^b	-6.35 ± 4.60 ^b	51.37 ± 5.44 ^b
<i>P</i> -value			0.0001 [*]	0.0020 [*]	0.0055 [*]	0.0001 [*]
Marker273509	AA	15	55.50 ± 4.41 ^a	0.85 ± 0.84 ^a	-8.82 ± 3.83 ^{ab}	45.69 ± 4.19 ^a
	GA	27	56.97 ± 3.60 ^a	1.15 ± 1.24 ^a	-9.30 ± 4.30 ^a	44.58 ± 3.32 ^a
	GG	239	50.77 ± 5.86 ^b	4.33 ± 2.63 ^b	-7.10 ± 4.47 ^b	50.97 ± 5.74 ^b
<i>P</i> -value			0.0001 [*]	0.0001 [*]	0.0235 [*]	0.0001 [*]

Table 5. Associations between three SNPs and shell nacre colour related traits in *H. cumingii*.

colour has always been an important criterion for evaluating pearl value, and shell colour is believed to have a connection with pearl colour because of their similarities in formation. A previous study revealed that pearl colour is related mainly to the shell nacre colour of the donor mussel in *H. cumingii*⁴⁸. The results showed that extremely significant correlation existed between inner shell color of the donor mussel and pearls color ($P < 0.01$). Moreover, the greater the dE (total colour change) value of the donor mussels is, the higher percent of purple pearl the mussel can produce. Therefore, pearl colour may be improved by improving shell nacre colour in *H. cumingii*.

A high-density linkage map is an effective tool for the fine mapping of QTLs for these important traits in *H. cumingii*. In this study, shell nacre colour and growth-related traits were fine mapped using the improved

linkage map, and 26 putatively significant QTLs were detected. Compared to our previous study, six more QTLs for growth-related traits were detected, and the positions of the QTLs were refined to a greater degree. However, the QTLs detected on LG1, LG4, LG8, and LG15 were significantly different from the results of our previous study³², in which growth-related QTLs were located on LG1, LG2, LG4, LG6, LG8, and LG18. Such inconsistent results are possibly due to the differences in the resolution of the two genetic maps. To date, one or two QTLs for some traditional growth-related traits, such as body weight, shell width, and shell weight, have been reported in many bivalves^{7,28,30,33,49}. In this study, seven QTLs for body weight, three QTLs for shell width, and two QTLs for shell weight were identified, with an average PVE of approximately 10%; thus, a higher QTL number and lower PVE were observed when compared with the abovementioned studies. In addition, five QTLs for inner mantle weight and two QTLs for margin mantle weight, which were considered to be related to pearl formation, were identified in our study. Since there is little doubt that shell formation and pearl formation are inextricably linked⁵⁰ and close ties exist between the growth traits and pearl traits, all the detected QTLs may have a potential impact on shell or pearl formation and should be investigated further. In the present study, seven QTLs associated with shell nacre colour were detected, showing a high percentage of PVE of 19.7% to 22.8%. Of these seven QTLs, five significant QTLs related to shell nacre colour tended to colocalize, clustering in a special region (17.3 cM to 36.1 cM) on LG17. This result showed a high similarity to the previously identified five QTLs located on the same chromosome in the region of 0–42.6 cM. In particular, the significant LOD region (17.3–36.1 cM) of these five QTLs contained 29 markers and showed smaller marker gaps of no more than 3.8 cM. These findings were more refined than those of our previous study in which only seven markers were found and the biggest gap was 16.2 cM (0–42.6 cM). Furthermore, three SNPs located on four QTLs were selected and validated to be significantly associated with shell nacre colour in other families, which suggested that the shell nacre colour related QTLs identified in the study had high accuracy and reliability. Although there are few studies on QTLs for shell nacre colour, a recent study on *P. maxima* reported nine putative QTLs for pearl colour and one QTL for pearl surface complexion³⁴. Moreover, the most prominent QTLs were located within a 2-cM interval on LG10, which is similar to our findings on QTLs for shell nacre colour. Given the correlation between shell nacre colour and pearl colour, we believe that there are some important genes or loci that determine shell nacre colour or pearl colour on LG17. Further studies need to be performed to identify genes that determine shell nacre colour in *H. cumingii*.

Methods

Mapping and DNA isolation. An F₁ full-sib family of *H. cumingii* produced by an intra-specific cross between two mussels was used as the mapping family. The female and male parents are from the Poyang Lake population and Dongting Lake population, respectively. The aquaculture environment and management process were as described by Bai *et al.*³². DNA was extracted from 200 F₁ individuals and their parents, according to a method described previously⁵¹. Finally, parentage analysis was carried out, and paternity of 157 F₁ progenies was confirmed using eight microsatellite markers, as described previously⁵².

SLAF library preparation and sequencing. The SLAF library was prepared as described previously²⁴, with minor modifications. A pre-designed experiment was conducted to evaluate the enzymes and restriction fragment sizes. On the basis of the results of the pilot experiment, the optimum scheme was confirmed and used to construct the SLAF library. Genomic DNA from the two parents and F₁ progenies was incubated at 37 °C with *MseI* (New England Biolabs, NEB), T4 DNA ligase (NEB), ATP (NEB), and *MseI* adapter. Restriction-ligation reactions were heat-inactivated at 65 °C, and digestion was performed using *HaeIII* and *PvuII-HFTM* at 37 °C. PCR was performed with a diluted restriction-ligation mixture, dNTP, *Taq* DNA polymerase (NEB), and *MseI* primer containing barcode 1. Next, we used the E.Z.N.A. Cycle Pure Kit (Omega) to purify and pool the PCR products. The pooled PCR products were subsequently incubated at 37 °C with *MseI*, T4 DNA ligase, ATP, and Solexa adapter. Then, the sample was purified using a Quick Spin column and run on a 2% agarose gel. Subsequently, fragments ranging from 314 to 394 bp were excised and purified using a Gel Extraction Kit, and these DNA fragments were amplified using PCR with the Phusion Master Mix (NEB) and Solexa Amplification primer mix to add barcode 2. Then, DNA fragments of 314–394 bp were gel-purified and diluted for pair-end sequencing on an Illumina High-seq 2500 sequencing platform at Biomarker Technologies Corporation in Beijing.

SLAF-seq data analysis and genotyping. SLAF-seq data analysis and genotyping were performed according to the method developed by Sun *et al.*²⁴. Briefly, all SLAF pair-end reads with clear index information were clustered on the basis of sequence similarity detected using BLAT⁵³. Then, sequences with more than 90% similarity were assigned to the same locus. Minor allele frequency (MAF) evaluation was subsequently performed to define alleles in each SLAF. *H. cumingii* is a diploid species, and one locus contains no more than four SLAF tags; therefore, loci containing more than four tags were discarded as repetitive SLAFs. Correspondingly, groups with two, three, and four tags were identified as polymorphic and considered potential SLAFs. To further improve SLAF accuracy, SLAFs with more than 80% integrity and more than 20-fold average sequence depths in parents were used for the map construction.

Construction of the genetic linkage map. To reduce the influence of distorted markers, the chi-square test (χ^2) was performed and the SLAF markers with significantly distorted segregation ($P < 0.05$) were excluded from the linkage mapping. Then, all high-quality SLAF markers and 506 SSR markers of the previous study were used for the linkage map construction³². The markers were divided into LGs by using the single-linkage clustering algorithm at logarithm of odds (LOD) threshold ≥ 4.0 and a maximum recombination fraction of 0.4. To avoid genotyping errors and deletion caused by NGS, High Map Strategy was used⁴⁵. MSTmap algorithms⁵⁴ and SMOOTH algorithms⁵⁵ were used to order the SLAF markers and correct genotyping errors, respectively. The LGs were constructed as follows: Primary marker orders were first obtained by their location on the chromosomes,

according to the relationship between ordered markers, and genotyping errors or deletions were corrected using the SMOOTH algorithm; then, MSTmap was used to order the map and SMOOTH was again used to correct the newly ordered genotypes. The processes were performed iteratively to ensure the accuracy of marker order and map distances, and high-quality maps were constructed after four or more cycles.

Phenotypic measurement and QTL analysis. The following five growth-related traits were measured in the 157 progenies: shell width, body weight, shell weight, margin mantle weight, and inner mantle weight and five shell colour-related traits, purple mantle scar length, CIE 1976 L* (lightness), a* (redness), b* (yellowness) colour space, and total colour change (dE)³². The average L* (AL), average a* (Aa), average b* (Ab), and average dE (AdE) represent average values measured on the three points of inner shell. Statistical analysis of the phenotypic data was conducted with SPSS 18.0. MapQTL 5.0 software was used for the QTL mapping with the interval mapping method. The genome was scanned at 1-cM intervals, the maximum LOD score along the interval was considered as the position of the QTL, and the region in the LOD score greater than the threshold was considered as the confidence interval. The LOD score threshold was initially set at 3.0 for QTL declaration, and QTLs that exceeded this LOD threshold were considered as suggestive QTLs. If any relevant QTL was identified, the LOD score threshold was determined using the 1,000-permutation test with a confidence of 0.99. QTLs with LOD scores greater than the threshold at a confidence of 0.99 were declared significant.

The validation of four QTLs associated with shell nacre colour on LG17. To examine whether the SNPs located on QTLs were associated with shell nacre colour, three SNPs were selected from shell nacre colour related QTLs (AL-1, Aa-1, Ab-3, AdE-1) and used for primer design (Table S5). Then, the primers and markers were assessed in three families of *H. cumingii* (one hundred individuals for each family) from Chongming aquaculture farm of Shanghai Ocean University in Shanghai, China. Genomic DNA was extracted and CIE 1976 L* (lightness), a* (redness), b* (yellowness) colour space, total colour change (dE) were measured as described above. Primer designing and genotyping assays were carried out using KASP technology (LGC Genomics, Hoddesdon, UK) as described previously³⁶. Associations between SNPs and shell nacre colour related traits were analyzed using Chi-square test in JMP 8.0 software.

References

- Li, J. & Li, Y. Aquaculture in China—Freshwater Pearl Culture. *World Aquaculture*. **40**, 60–62 (2009).
- Massault, C., Bovenhuis, H., Haley, C. & de Koning, D.-J. QTL mapping designs for aquaculture. *Aquaculture*. **285**, 23–29 (2008).
- Wang, X.-j., Wang, Y.-j., Shi, A.-j. & Wang, X.-z. Research on chromosomes of *Hyriopsis cumingii*. *Sichuan Daxue Xuebao (Ziran Kexueban)* **37**, 252–256 (2000).
- Wringe, B. F. *et al.* Growth-related quantitative trait loci in domestic and wild rainbow trout (*Oncorhynchus mykiss*). *BMC Genet.* **11**, 63 (2010).
- Wang, L. *et al.* Construction of a high-density linkage map and fine mapping of QTL for growth in Asian seabass. *Sci Rep.* **5**, 16358 (2015).
- Yue, G. H. Recent advances of genome mapping and marker-assisted selection in aquaculture. *Fish. Fish.* **15**, 376–396 (2014).
- Jiao, W. *et al.* High-Resolution Linkage and Quantitative Trait Locus Mapping Aided by Genome Survey Sequencing: Building Up An Integrative Genomic Framework for a Bivalve Mollusc. *DNA Res.* **21**, 85–101 (2014).
- Liu, Z. J. & Cordes, J. F. DNA marker technologies and their applications in aquaculture genetics. *Aquaculture* **238**, 1–37 (2004).
- Leonforte, A. *et al.* SNP marker discovery, linkage map construction and identification of QTLs for enhanced salinity tolerance in field pea (*Pisum sativum* L.). *BMC. Plant. Biol.* **13**, 14 (2013).
- Perkel, J. SNP genotyping: six technologies that keyed a revolution. *Nat. Methods* **5**, 575 (2008).
- Davey, J. W. & Blaxter, M. L. RADSeq: next-generation population genetics. *Brief. Funct. Genomics* **10**, 108 (2011).
- Wang, S., Meyer, E., McKay, J. K. & Matz, M. V. 2b-RAD: a simple and flexible method for genome-wide genotyping. *Nat Methods* **9**, 808–810 (2012).
- Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S. & Hoekstra, H. E. Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *Plos One* **7**, e37135 (2012).
- Elshire, R. J. *et al.* A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *Plos One* **6**, e19379 (2011).
- Houston, R. D. *et al.* Development and validation of a high density SNP genotyping array for Atlantic salmon (*Salmo salar*). *BMC Genomics* **15**, 13 (2014).
- Liu, S. *et al.* Development of the catfish 250K SNP array for genome-wide association studies. *BMC. Res. Notes* **7**, 135 (2014).
- Xu, J. *et al.* Development and evaluation of the first high-throughput SNP array for common carp (*Cyprinus carpio*). *BMC Genomics* **15**, 307 (2014).
- Pocwierz-Kotus, A. *et al.* Genetic differentiation of brackish water populations of cod *Gadus morhua* in the southern Baltic, inferred from genotyping using SNP-arrays. *Mar Genomics* **19**, 17–22 (2015).
- Palti, Y. *et al.* The development and characterization of a 57K single nucleotide polymorphism array for rainbow trout. *Mol. Ecol. Resour.* **15**, 662–672 (2015).
- Gonen, S. *et al.* Linkage maps of the Atlantic salmon (*Salmo salar*) genome derived from RAD sequencing. *BMC Genomics* **15**, 166 (2014).
- Amish, S. J. *et al.* RAD sequencing yields a high success rate for westslope cutthroat and rainbow trout species-diagnostic SNP assays. *Mol. Ecol. Resour.* **12**, 653–660 (2012).
- Li, Y. *et al.* Construction of a high-density, high-resolution genetic map and its integration with BAC-based physical map in channel catfish. *DNA Res.* **22**, 39–52 (2015).
- Shao, C. *et al.* Genome-wide SNP identification for the construction of a high-resolution genetic map of Japanese flounder (*Paralichthys olivaceus*): applications to QTL mapping of *Vibrio anguillarum* disease resistance and comparative genomic analysis. *DNA Res.* **22**, 161–170 (2015).
- Sun, X. *et al.* SLAF-seq: an efficient method of large-scale *de novo* SNP discovery and genotyping using high-throughput sequencing. *PLoS One* **8**, e58700 (2013).
- Xu, X. *et al.* A high-density genetic map of cucumber derived from Specific Length Amplified Fragment sequencing (SLAF-seq). *Front. Plant. Sci.* **5**, 768 (2014).
- Qi, Z. *et al.* A high-density genetic map for soybean based on specific length amplified fragment sequencing. *Plos One* **9**, e104871 (2014).
- Ma, J. Q. *et al.* Large-Scale SNP Discovery and Genotyping for Constructing a High-Density Genetic Map of Tea Plant Using Specific-Locus Amplified Fragment Sequencing (SLAF-seq). *Plos One* **10**, e0128798 (2015).

28. Guo, X., Li, Q., Wang, Q. Z. & Kong, L. F. Genetic Mapping and QTL Analysis of Growth-Related Traits in the Pacific Oyster. *Mar. Biotechnol.* **14**, 218–226 (2011).
29. Sauvage, C. *et al.* QTL for resistance to summer mortality and OsHV-1 load in the Pacific oyster (*Crassostrea gigas*). *Anim. Genet.* **41**, 390–399 (2010).
30. Li, H., Liu, X. & Zhang, G. A consensus microsatellite-based linkage map for the hermaphroditic bay scallop (*Argopecten irradians*) and its application in size-related QTL analysis. *Plos One* **7**, e46926 (2012).
31. Qin, Y., Liu, X., Zhang, H., Zhang, G. & Guo, X. Identification and mapping of amplified fragment length polymorphism markers linked to shell color in bay scallop, *Argopecten irradians* (Lamarck, 1819). *Mar Biotechnol (NY)* **9**, 66–73 (2007).
32. Bai, Z. *et al.* Constructing a microsatellite-based linkage map and identifying QTL for pearl quality traits in triangle pearl mussel (*Hyriopsis cumingii*). *Aquaculture* **437**, 102–110 (2015).
33. Jones, D. B. *et al.* Determining genetic contributions to host oyster shell growth: Quantitative trait loci and genetic association analysis for the silver-lipped pearl oyster, *Pinctada maxima*. *Aquaculture* **434**, 367–375 (2014).
34. Jones, D. B. *et al.* Quantitative trait loci and genetic association analysis reveals insights into complex pearl quality traits in donor silver-lipped pearl oysters. *Aquaculture* **434**, 476–485 (2014).
35. Kujur, A. *et al.* Ultra-high density intra-specific genetic linkage maps accelerate identification of functionally relevant molecular tags governing important agronomic traits in chickpea. *Sci Rep.* **5**, 9468 (2015).
36. Tamura, Y. *et al.* Map-based cloning and characterization of a brown planthopper resistance gene *BPH26* from *Oryza sativa* L. ssp. *indica* cultivar ADR52. *Sci Rep.* **4**, 5872 (2014).
37. Lee, M. *et al.* A consensus linkage map of oil palm and a major QTL for stem height. *Sci Rep.* **5**, 8232 (2015).
38. Hyten, D. L. *et al.* High-throughput SNP discovery through deep resequencing of a reduced representation library to anchor and orient scaffolds in the soybean whole genome sequence. *BMC Genomics* **11**, 38 (2010).
39. Brown, P. O. & Botstein, D. Exploring the new world of the genome with DNA microarrays. *Nat. Genet.* **21**, 33–37 (1999).
40. Li, B. *et al.* Construction of a high-density genetic map based on large-scale markers developed by specific length amplified fragment sequencing (SLAF-seq) and its application to QTL analysis for isoflavone content in *Glycine max*. *BMC Genomics* **15**, 1086 (2014).
41. Cai, C., Cheng, F. Y., Wu, J., Zhong, Y. & Liu, G. The First High-Density Genetic Map Construction in Tree Peony (*Paeonia Sect. Moutan*) using Genotyping by Specific-Locus Amplified Fragment Sequencing. *Plos One* **10**, e0128584 (2015).
42. 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* **14**, 810 (2013).
43. Shi, Y. *et al.* High-density single nucleotide polymorphisms linkage and quantitative trait locus mapping of the pearl oyster, *Pinctada fucata martensii* Dunker. *Aquaculture* **434**, 376–384 (2014).
44. Ferreira, A., da Silva, M. F., Silva, L. & Cruz, C. D. Estimating the effects of population size and type on the accuracy of genetic maps. *Genet. Mol. Biol.* **29**, 187–192 (2006).
45. Liu, D. *et al.* Construction and analysis of high-density linkage map using high-throughput sequencing data. *PLoS One* **9**, e98855 (2014).
46. Bai, Z., Li, J. & Wang, G. Relationship between pearl production, growth traits and the inserted position of mantle piece in triangle mussel (*Hyriopsis cumingii*). *J. Fish. Sci. China* **15**, 493–499 (In Chinese with English abstract) (2008).
47. Li, J. & Liu, Y. The main influencing factors on the quality of cultured pearls. *J. fish. China* **35**, 1753–1760 (In Chinese with English abstract) (2011).
48. Li, Q. *et al.* Correlation analysis of non-nucleated pearl quality parameters with growth traits and inner shell color of *Hyriopsis cumingii*. *J. Fish. Sci. China* **39**, 1632–1639 (In Chinese with English abstract) (2015).
49. Zhan, A. *et al.* Construction of microsatellite-based linkage maps and identification of size-related quantitative trait loci for Zhikong scallop (*Chlamys farreri*). *Anim. Genet.* **40**, 821–831 (2009).
50. Suzuki, M. *et al.* Microstructure of iridescence-lacking pearl formed in *Pinctada fucata*. *J. Cryst. Growth* **433**, 148–152 (2016).
51. Yue, G. H. & Orban, L. A simple and affordable method for high-throughput DNA extraction from animal tissues for polymerase chain reaction. *Electrophoresis* **26**, 3081–3083 (2005).
52. Bai, Z., Zheng, H., Lin, J., Wang, G. & Li, J. Comparative analysis of the transcriptome in tissues secreting purple and white nacre in the pearl mussel *Hyriopsis cumingii*. *PLoS One* **8**, e53617 (2013).
53. Doerks, T., Copley, R. R., Schultz, J., Ponting, C. P. & Bork, P. Systematic identification of novel protein domain families associated with nuclear functions. *Genome Res.* **12**, 47–56 (2002).
54. Wu, Y., Bhat, P. R., Close, T. J. & Lonardi, S. Efficient and accurate construction of genetic linkage maps from the minimum spanning tree of a graph. *PLoS Genet* **4**, e1000212 (2008).
55. van Os, H., Stam, P., Visser, R. G. & van Eck, H. J. SMOOTH: a statistical method for successful removal of genotyping errors from high-density genetic linkage data. *Theor Appl Genet* **112**, 187–194 (2005).
56. Chopra, R. *et al.* Next-generation transcriptome sequencing, SNP discovery and validation in four market classes of peanut, *Arachis hypogaea* L. *Mol. Genet. Genomics* **290**, 1169–1180 (2015).

Acknowledgements

This work was supported by grants from the National Science and Technology Support Program (2012BAD26B04), the National Natural Science Foundation of China (31272657), Shanghai Collaborative Innovation Center for Aquatic Animal Genetics and Breeding (ZF1206), and the Project of Shanghai Engineering and Technology Center for Promoting Ability (13DZ2280500).

Author Contributions

J.-L.L. supervised the project. Z.-Y.B., X.-K.H. and X.-J.L. conceived and designed the study, guided data analysis and interpretation. X.-K.H. performed experiments and drafted the manuscript. Q.-Q.L. and X.-K.H. collected samples. All authors approved the final version of the manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Bai, Z.-Y. *et al.* Construction of a high-density genetic map and QTL mapping for pearl quality-related traits in *Hyriopsis cumingii*. *Sci. Rep.* **6**, 32608; doi: 10.1038/srep32608 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016