OPEN ACCESS International Journal of Molecular Sciences ISSN 1422-0067 www.mdpi.com/journal/ijms

Short Note

Development of 101 Gene-based Single Nucleotide Polymorphism Markers in Sea Cucumber, *Apostichopus japonicus*

Huixia Du, Zhenmin Bao, Jingjing Yan, Meilin Tian, Xiaoyu Mu, Shi Wang and Wei Lu *

Key Laboratory of Marine Genetics and Breeding, College of Marine Life Science, Ocean University of China, Qingdao 266003, China; E-Mails: nevergiveupxia@sina.com (H.D.); zmbao@ouc.edu.cn (Z.B.); jing-yan12345@163.com (J.Y.); air880426@163.com (M.T.); muxiaoyu422@yahoo.cn (X.M.); swang@ouc.edu.cn (S.W.)

* Author to whom correspondence should be addressed; E-Mail: lw1981@ouc.edu.cn; Tel./Fax: +86-532-8203-1802.

Received: 25 April 2012; in revised form: 25 May 2012 / Accepted: 25 May 2012 / Published: 8 June 2012

Abstract: Single nucleotide polymorphisms (SNPs) are currently the marker of choice in a variety of genetic studies. Using the high resolution melting (HRM) genotyping approach, 101 gene-based SNP markers were developed for *Apostichopus japonicus*, a sea cucumber species with economic significance for the aquaculture industry in East Asian countries. HRM analysis revealed that all the loci showed polymorphisms when evaluated using 40 *A. japonicus* individuals collected from a natural population. The minor allele frequency ranged from 0.035 to 0.489. The observed and expected heterozygosities ranged from 0.050 to 0.833 and 0.073 to 0.907, respectively. Thirteen loci were found to depart significantly from Hardy–Weinberg equilibrium (HWE) after Bonferroni corrections. Significant linkage disequilibrium (LD) was detected in one pair of markers. These SNP markers are expected to be useful for future quantitative trait loci (QTL) analysis, and to facilitate marker-assisted selection (MAS) in *A. japonicus*.

Keywords: single nucleotide polymorphism (SNP); *Apostichopus japonicus*; high resolution melting (HRM) analysis; marker-assisted selection (MAS)

1. Introduction

The sea cucumber *Apostichopus japonicus* (Selenka 1867), naturally distributes along the coasts of China, Japan, Korea and Russia [1]. Due to their nutritional and medicinal value, they have long been exploited as an important fishery resource in East Asian countries. Over the past decade, the aquaculture of *A. japonicus* has become widespread along the coasts of China, due to increasing market demand and over-exploitation of wild sea cucumbers [2]. However, the rapid expansion and intensification of sea cucumber aquaculture has resulted in some severe problems, such as wide-spread disease and stock deterioration, possibly caused by inappropriate broodstock management and inbreeding depression [2]. In order to properly manage broodstock resources and efficiently enhance aquaculture production, control of inbreeding and selection of broodstock with the desired traits, such as rapid growth and disease resistance, are currently necessary for sustainable development of the *A. japonicus* aquaculture. Recently, marker-assisted selection (MAS) has become a valuable tool for selecting individuals with traits of interest [3]. To perform MAS, a large number of genetic markers are usually needed to determine the quantitative trait loci (QTLs) associated with economically important traits.

Single nucleotide polymorphisms (SNPs) have been shown to be the most abundant type of genetic variations in eukaryotic genomes [4], and are currently the marker of choice in a variety of genetic studies, such as high-density genetic linkage mapping and QTL analysis. However, only a limited number of SNP markers have been reported for *A. japonicus* [5–7]. Moreover, molecular markers developed from the expressed sequence tag (EST) databases offer several advantages over anonymous genomic markers, as (i) they can detect variation in the expressed portion of the genome, so that gene tagging could give "perfect" marker-trait associations; (ii) they could alleviate the problem of null alleles which is usually associated with markers developed from the non-transcribed regions; and (iii) they are expected to have greater transferability between species, since transcribed regions are more conserved among closely related species/genera.

Previously, our group has released a large amount of EST data by 454 sequencing of the *A. japonicus* transcriptome [7]. By mining our EST dataset, more than 54,000 putative SNPs have been identified, 200 of which were selected in this study for marker development. SNP validation was performed using 48 *A. japonicus* individuals collected from four natural populations. Genetic parameters of the validated SNP markers were evaluated using 40 *A. japonicus* individuals from a single natural population. These SNP markers will be useful for future QTL analysis in order to facilitate MAS in *A. japonicus*.

2. Results and Discussion

Transcriptomic sequences represent an important resource for rapid and cost-effective development of gene-based SNPs. For the high resolution melting (HRM)-based SNP marker development, we designed PCR primers for 200 candidate SNPs (Table 1), which were previously identified from the *A. japonicus* transcriptome generated by 454-FLX sequencing [7]. After PCR amplification, 159 (79.5%) amplified strong bands with expected sizes. The others were discarded without further consideration, as they produced bands larger than expected (possibly caused by introns) or resulted in poor amplification (weak or non-specific amplification). During the initial HRM screen, 63.5% (101) of the 159 successfully

amplified loci showed polymorphisms in 48 individuals collected from 4 natural populations, 21.4% (34) generated non-polymorphic curves, and 15.1% (24) displayed unreliable melting curves. In this study, we showed that minor allele frequency (MAF) can serve as an important selection criterion to distinguish true SNPs from sequencing errors when performing SNP mining from 454 sequencing data (Figure 1). For example, most of the validated SNPs usually have a MAF of more than 35%, whereas most non-validated SNPs usually have a MAF of less than 25%. Although our study demonstrated that SNP markers can be efficiently developed from transcriptomic resources, it should be noted that the SNPs obtained may largely represent common genetic variations due to the low coverage of the original transcriptome sequencing, and may suffer from ascertainment bias resulting from simple sample source used in the original transcriptome sequencing.

 Table 1. Results of validation and genotyping of candidate single nucleotide polymorphisms (SNPs).

Categories	Number of SNPs
Total number of tested SNPs	200
Successful PCR	159
Successful genotype calling	135
Polymorphic SNPs	101
Monomorphic SNPs	34
Failed SNPs	65

Figure 1. Distribution of SNP minor allele frequency (MAF) for *Apostichopus japonicus*. The number above each bar was the polymorphic rate in respective MAF categories.



Genetic parameters of the validated SNP markers were further evaluated using 40 *A. japonicus* individuals from a single natural population. As expected, all 101 SNP loci were polymorphic. The minor allele frequency ranged from 0.035 to 0.489 (Table 2). The *H*o ranged from 0.050 to 0.833, while the *H*e varied from 0.073 to 0.907. Thirteen loci departed significantly (p < 0.01) from Hardy–Weinberg equilibrium (HWE) after Bonferroni correction, suggesting that these loci may be under ongoing natural selection. Significant linkage disequilibrium (LD) was detected in one pair of SNP markers (ApjSNP092_CT and ApjSNP098_CT).

Locus ID	Gene Name	Primers and Probes (5'–3')	Size (bp)	Ho	He	MA	MAF	<i>p</i> -Value
ApjSNP001_CT	similar to Mech2 protein	F:CCTCAGTCCCAATCACCACT	98	0.250	0.431	Т	0.128	0.239
		R:ACACTGGCATACACCAGCAA						
		P:CAATGACTTCTCCTTCCTACAGCTTCC						
ApjSNP002_CT	Iron-sulfur cluster assembly 2 homolog	F:TGAATCAGGCAGTTGTGATGA	102	0.231	0.485	С	0.397	0.080
		R:GGTCCAGCTAGTCATGCTTTT						
		P:TCAAAGAAATACACTA <u>T</u> TTCATCCGATAAAGCAAG						
ApjSNP003_AC	Protein strawberry notch homolog 2	F:AGCGATTATATCCGATGCAG	108	0.431	0.583	А	0.489	0.482
		R:GCTGACCAGGGTAGTTCGAC						
		P:TACAAGAGTAGCCAATCAGAGGCGAGC						
ApjSNP004_AG	Thiosulfate sulfurtransferase	F:CAGTTGTAACTGCACCTCAGC	70	0.314	0.342	А	0.178	0.578
		R:ATGCCTACTTGGATGCCAGA						
		P:ACCACAGGGTGT <u>A</u> GCCAGGTTCGTCAG						
ApjSNP005_AG	Thiosulfate sulfurtransferase	F:AGGCATCCCTACGGGTATTT	70	0.374	0.312	А	0.240	0.857
		R:ACAGGAATGAAGTGGCTTGG						
		P:TCGCTCTAGTCATCCGCCTCAACG						
ApjSNP006_CT	Dynein heavy chain 6, axonemal	F:GGGAGGTCTTACGAAGTGGA	97	0.271	0.273	С	0.384	0.345
		R:CGAGGAGTTCGGAGTAGCTTT						
		P:ACCCTACCGACTGTCGCAGAGAATG						
ApjSNP007_AG	Sodium-dependent phosphate transport protein 2B	F:ACCTTGGTGGCAGATATGGA	75	0.252	0.432	G	0.287	0.418
		R:TTCAGTGTCCGCAGATTTCTT						
		P:AGTTGAATTAACGGCTCTCGAACCAG						
ApjSNP008_GT	Testis-specific serine/threonine-protein kinase 1	F:CCACAATTAGCGATGGGTTT	102	0.349	0.488	G	0.410	0.058
		R:CAAAGCTCCAGGACTTCTGC						
		P:GTTGTAGTGAACGCA <u>T</u> TGGTTTAGGAAGGAAAC						
ApjSNP009_AC	Disintegrin and metalloproteinase domain-containing	F:CTAAAGGGGATCACCACGAC	94	0.208	0.289	А	0.104	0.365
		R:ATAAGCAGGCTTCCCTTTCG						
		P:ACCTTTCGCCGGCCACGCCCT						

Table 2. Characterization of 101 SNPs for the sea cucumber Apostichopus japonicus.

Table 2. Cont.

Locus ID	Gene Name	Primers and Probes (5'–3')	Size (bp)	Ho	He	MA	MAF	<i>p</i> -Value
ApjSNP010_GT	Zinc finger protein 62 homolog	F:CCACCAGATGTCTTTGATTCG	106	0.428	0.512	G	0.448	0.552
		R:TCACGACCAATACTGCTTGG						
		P:AGATCCGACCCATGCAAGACCAAGGT						
ApjSNP011_AG	Kelch-like protein 9	F:CAGTCAGCCTAGCCCTACCA	93	0.500	0.498	G	0.458	0.574
		R:TCGTTGACCTTTGGTACTGATG						
		P:GTGCAAACCAATC <u>G</u> CAAGTCATTGTCGT						
ApjSNP012_CT	TATA box-binding protein-associated factor RNA	F:CCTTCACTGGTATGGCATGTT	88	0.314	0.468	Т	0.240	0.045
		R:TGATCCATGTAGGGAGGCTTT						
		P:GTACATTAACT <u>C</u> TCCACAAGCTCCCTTGTA						
ApjSNP013_AT	Protocadherin Fat 3	F:TGTTAGCACCTCTATCAAGGATGA	102	0.500	0.454	А	0.454	1.000
		R:TTCCATACCTCCTGCCAATC						
		P:GTTCAAGGACACTTG <u>A</u> TGGAAAGTGTAATGATT						
ApjSNP014_GT	Seryl-tRNA synthetase, mitochondrial	F:ATTCGTGTCCAGTTCGCAAT	96	0.271	0.276	G	0.386	0.346
		R:GAGATCGGGCGATATAACCA						
		P:TCATATCAATTTGT <u>G</u> CCTCGAGGATCGAC						
ApjSNP015_CT	Creatine kinase, flagellar	F:TCACAGGCCATCGATCATAC	92	0.436	0.502	С	0.446	0.556
		R:CCTTTTCACCAACCTCTCCA						
		P:TCTAAGAGGTG <u>C</u> TGGTGCCCAGTAC						
ApjSNP016_AG	Fibrinogen-like protein A	F:AATGGCCTCAAGAAAGTGGA	108	0.430	0.583	А	0.483	0.497
		R:TCCAGTACCTAGATTTGAAGGACA						
		P:GAATTCATGTGGA <u>G</u> TGAGCATCTTGGAAT						
ApjSNP017_GT	Abhydrolase domain-containing protein 14B	F:CGGGGTCTACCTCATACAACC	78	0.293	0.444	Т	0.475	0.854
		R:CCTCCGCCATCTACAGTGTT						
		P:CATATATGGAGCCATT <u>T</u> GCTGTATATTGTAACATG						
ApjSNP018_AG	Apolipoprotein A–I-binding protein	F:CATAGGTGTCCAGAAATGTTCG	93	0.073	0.083	G	0.083	1.000
		R:TGTCCCATGTCTAAAGCATAACTG						
		P:CACAGAGTTCCCATGGGCAGATAGAAG						

Locus ID	Gene Name	Primers and Probes (5'-3')	Size (bp)	Ho	He	MA	MAF	<i>p</i> -Value
ApjSNP019_AG	N-acyl-phosphatidylethanolamine-hydrolyzing	F:CGTGCTCGGTTTTAATGTTG	91	0.688	0.505	А	0.354	0.498
		R:CATGGTGAAACCTGGTAGACG						
		P:CCAAGCACAACCAGAAACCGAGAAATCCA						
ApjSNP020_AT	Polypeptide N-acetylgalactosaminyltransferase 11	F:AAAGAGGTATCGACCTTGTCCA	109	0.250	0.256	А	0.328	0.857
		R:TGCTCGGACTGTATGTTCATC						
		P:TGGAGGAACTTCC <u>A</u> GAAATCAATGCTGAG						
ApjSNP021_AG	Hyalin	F:TTCAAGTGGTATCACGAAAACG	92	0.108	0.333	А	0.290	0.557
		R:CGTGCTATTGCCTTTGGATT						
		P:GCTGAGGCTTCC <u>A</u> AAAGATGACGATTC						
ApjSNP022_CG	Transmembrane protein 129	F:TGGAATGCCACTAACACCAA	80	0.442	0.364	G	0.483	0.381
		R :TTGACACCACCACCAATC						
		P:TTGATATGTCTG <u>C</u> TGGGCTATTCTGGTA						
ApjSNP023_CT	Mediator of RNA polymerase II transcription subunit	F:GCTGATGAGCAATCTTCACACT	95	0.146	0.505	Т	0.489	0.051
		R:CAAGTTTCAGACGGGACCTG						
		P:GTCTTGATTATCCACGAATCTGTGACATACCA						
ApjSNP024_AG	AF339450_1 hillari	F:TCCATTGAACGGAGGACTTC	108	0.419	0.484	А	0.395	0.376
		R:CAAACATTTCAGCCTTGTGG						
		P:GTCTGGGATGG <u>G</u> ATGTAGTCGACACTTA						
ApjSNP025_AC	Proteasome subunit beta type-5	F:TCCAGATCGCTACGGTCTTC	81	0.250	0.250	А	0.423	0.125
		R:ACGACCAGGTAGCTGCAGAG						
		P:TGGTGTATCAAGGA <u>A</u> ATTCAAACCCAGCTGT						
ApjSNP026_AG	Dynactin subunit 5	F:GCCTGTTGCTGTTAACTTTCG	110	0.316	0.365	А	0.461	0.724
		R:CTGGCATGTAACTCTATGAAACTC						
		P:GTTAAGTGAAAGTTGACTGCCTCAGTATTGTA						
ApjSNP027_AG	Apoptosis-inducing factor 2	F:CAGAGAAAGCTGGAGATGATGA	88	0.516	0.467	А	0.361	0.324
		R:ATGATTTCAACTGGGCCATC						
		P:GATGATGAACC <u>G</u> CAGAAGGGTTCGAA						

Table 2. Cont.

Locus ID	Gene Name	Primers and Probes (5'–3')	Size (bp)	Ho	He	MA	MAF	<i>p</i> -Value
ApjSNP028_CT	Uncharacterized protein C6orf163	F:ATAGTTGGGTGTGGCTTTGC	104	0.209	0.190	С	0.105	0.698
		R:CCGATGCAGTGATGGAAATA						
		P:AAATGTCACCTAACTGTGATTGATCCTCGCC						
ApjSNP029_AT	F-box/LRR-repeat protein 2	F:CCGTGATCCTAAATGAGGCTA	98	0.271	0.237	А	0.135	0.762
		R:CGCTAAGAGTAAGAGAAAGAAGCA						
		P:GCCTAACCATACTGGATTGGCTAGCAGT						
ApjSNP030_CT	TBC1 domain family member 10B	F:CCGGAGACGTAAAAGCACTC	91	0.191	0.174	Т	0.095	0.754
		R:TCGTCGTGTCTGGTATCCAC						
		P:AAGTCTGGACAG <u>C</u> TGTTAGCTAAGGGC						
ApjSNP031_CG	Stejaggregin-A subunit alpha	F:ATCGGTGCTAGACCCAAAGA	81	0.150	0.245	G	0.264	0.358
		R:TCCTTCTCTGGTGAATTGATTG						
		P:CATCCCAACGACGACGACCGATATGGTA						
ApjSNP032_AC	Lysine-specific demethylase 6A	F:CGAAGGCAACCAAGTAGGAC	91	0.138	0.833	С	0.383	0.497
		R:TGCCACCTCGATCATTTTCT						
		P:CGCTGGTGTTAATAACTTCATAGTCCGTTAC						
ApjSNP033_AG	ATP synthase subunit beta, mitochondrial	F:GAGTAACAACGGCCCAGAAA	76	0.458	0.467	А	0.232	0.854
		R:TACAGTGCCTACACCGGTCA						
		P:GGTCTGACCGCT <u>A</u> TTGGGATCAATCTGC						
ApjSNP034_GT	Ubiquitin carboxyl-terminal hydrolase 8	F:GGCTTGAAGAAACATGGGTAA	110	0.292	0.314	G	0.035	1.000
		R:CCAGTAGATTGCATCTTTCCATC						
		P:TCATGTTCACTTCTT <u>T</u> ATACCACACGATGACAT						
ApjSNP035_CT	Uncharacterized protein C7orf26 homolog	F:CGGTGGTGAGGTGTCTACATT	76	0.449	0.367	Т	0.485	0.498
		R:GGAATAGGCAACTCGAGGAA						
		P:GTCGGTGAAGTA <u>C</u> GAAGCCTTCATGAA						
ApjSNP036_AC	hypothetical protein	F:AAGATGCCAGACAGCAACAA	100	0.545	0.413	С	0.264	0.857
		R:CATGACTGCGTCTTCTGCTC						
		P:CAGGAATCTCAC <u>A</u> GACGAGAGGGAACT						

Table 2. Cont.

Locus ID	Gene Name	Primers and Probes (5'-3')	Size (bp)	Ho	He	MA	MAF	<i>p</i> -Value
ApjSNP037_AG	DNA replication licensing factor MCM8	F:GGAACCGGAGAGATGACAGA	95	0.492	0.502	А	0.458	0.557
		R:CCAGCGTCGTCACCTTTTAC						
		P:AGAGCAAGATCAAC <u>A</u> GAATGAGGACAAAGTA						
ApjSNP038_AG	LRP2-binding protein	F:GATGAAAGTACCTGGGAGGAA	83	0.750	0.625	G	0.147	0.381
		R:AGCTGATCATCGGTCCATCT						
		P:GGAGATTGAAGATT <u>G</u> ATCCCACTGACAAACTC						
ApjSNP039_AG	Endoplasmin	F:ATAACGTCGGACGAGCATTC	76	0.409	0.479	G	0.387	0.051
		R:AGCAACCACCATCTCTCTGC						
		P:AAGGGTTTGGAGT <u>A</u> AAACAGTCGGATGCCC						
ApjSNP040_CT	heat shock protein 90 kDa beta	F:CTTTGAAGATATGATGCCCAAG	102	0.348	0.291	С	0.174	0.084
		R:TTGTGTTGCTGCAGGGTTT						
		P:ACTCCGATGACCTGCCTCTCAATGTGA						
ApjSNP041_CT	Titin	F:AGCCATCGAGAATGAGAAGC	82	0.382	0.314	Т	0.192	0.091
		R:TGATGGTCTGTTCGATCCAC						
		P:GGTCACCGACTACGACAAGATCTCCTGC						
ApjSNP042_AG	Midasin	F:CAGCCTGGAAGACCCTCAGT	88	0.800	0.691	G	0.291	0.635
		R:TTGGACTTCCACCATCAGAA						
		P:AACCAGGCTAC <u>G</u> ATTTCATGGACCGGT						
ApjSNP043_CT	Scavenger receptor cysteine-rich type 1 protein M130	F:GGTTCACAACCTCAGGATGAC	95	0.317	0.505	С	0.476	0.200
		R:CTTCTGCACACCGCACTTT						
		P:GAAATTACAACCTGCTTTAGTGTCCAGAGATAG						
ApjSNP044_AC	FK506-binding protein 15	F:TCATACACTCAGGGCATCCA	90	0.583	0.473	А	0.332	0.149
		R:GCGTAGGCATATGACGAGAGA						
		P:CAGTTTTGTGAGTGT <u>C</u> TTGACAGTGATAGTGG						
ApjSNP045_AC	Titin	F:CGTTGAGATCCAAGTCAATGAG	105	0.512	0.502	А	0.456	0.897
		R:TGTAGGTGAGTGGTGAACGTG						
		P:TAGAAAGAATGGA <u>C</u> AGCGTCCCTGGAGT						

Locus ID	Gene Name	Primers and Probes (5'–3')	Size (bp)	Ho	He	MA	MAF	<i>p</i> -Value
ApjSNP046_AG	Radial spoke head protein 4 homolog A	F:GGGGAAGATGAGGTAGAAACG	81	0.113	0.109	G	0.056	0.623
		R:GCTCATACCGATTCCTGCTT						
		P:ACTCCCAAACCTACCGGAACTTATGTTTTAGA						
ApjSNP047_CT	Phenylalanyl-tRNA synthetase beta chain	F:TGGCAAATCAATCGGATTCT	102	0.326	0.300	Т	0.178	0.653
		R:AACGGTTCAATGGTTATCTCTAGG						
		P:CTCAAAGTTTGAGCTTCCAAACCCATGTGGA						
ApjSNP048_AG	Mitochondrial inner membrane protein	F:CCGATGAGAGGGGTATTCAA	98	0.222	0.468	А	0.361	0.002 *
		R:CCCCCATTCTCGTCTATCAG						
		P:GGGAGAGGTGGGA <u>G</u> AATATCCAGAGATA						
ApjSNP049_CT	Sulfotransferase family cytosolic 1B member 1	F:CCAGGGTAAAGTCAAAGGTCA	82	0.524	0.479	Т	0.278	0.401
		R:ACTGTAGCCCAGAACGATGC						
		P:TCCTTTCATTTTCCCCCTCGTACAAGTCATGT						
ApjSNP050_CT	RalA-binding protein 1	F:GGTTGAGGAGTTCTTGGGAGT	105	0.250	0.408	С	0.275	0.018
		R:CATCAGCATGATCCAACACA						
		P:CTGAATGATTTGCCAACTTGTAACTACACCTTAGA						
ApjSNP051_GT	Alpha-amylase B	F:TTCGATTCATCTGGTGCTTG	107	0.096	0.481	Т	0.390	0.005 *
		R:CTTGACCTTCGCAGGTGTTT						
		P:TGGAGAGAGATCCG <u>T</u> AACATGGTCGAATTGT						
ApjSNP052_GT	Putative vitellogenin receptor	F:CAGTCTGAAAGAACCACTGAAGA	98	0.411	0.485	G	0.400	0.758
		R:CGAGTATAGGAGGCTGAAAACG						
		P:GCCCAGAAGATATCGCCTCTCTTCAAATAGG						
ApjSNP053_CT	UDP-N-acetylglucosaminepeptide	F:TCGAAGCTAGATTACTGTGAGCA	101	0.071	0.503	Т	0.404	0.000 *
		R:TCTGAAGGAGATGCAGGACA						
		P:TGATTTGGATGGC <u>T</u> CTGGTATAGCACTCA						
ApjSNP054_CT	Kanadaptin	F:CAAGCCGTACATGAAAGCAA	88	0.585	0.506	С	0.489	0.307
		R:TGTCCAGGTACGAGTCATCG						
		P:AGAAGAAGAAGAA <u>T</u> TGGGCGGACGATCT						

Table 2. Cont.

Locus ID	Gene Name	Primers and Probes (5'–3')	Size (bp)	Ho	He	MA	MAF	<i>p</i> -Value
ApjSNP055_GT	Epidermal growth factor receptor	F:TCACGTTCCACCAGATTTTG	104	0.253	0.435	G	0.424	0.518
		R:ATGATGGGGGTAATGGCATA						
		P:TGACCAATAGCATATTCGATGTGATGTCACCA						
ApjSNP056_CT	hypothetical protein	F:ATGCCACCCTCTTAATCTGG	107	0.125	0.117	Т	0.318	0.442
		R:CTTGCCTGGGTTTTCCATAC						
		P:TCAGACCGGTGC <u>T</u> TCTGACAGTACATT						
ApjSNP057_CG	RuvB-like 2	F:CCATAACACCGATGACACCA	108	0.295	0.388	С	0.258	0.159
		R:GAAGCTGATAAGATGGAAGTAGCC						
		P:CATTGTCAAGGCA <u>G</u> TCATCTTGTCAGGA						
ApjSNP058_AG	Eyes absent homolog 1	F:CGTATCCCGTACCACAACCT	79	0.400	0.501	G	0.247	0.485
		R:AACCCGTAGGGAACCTGACT						
		P:GGTGTGCAACCA <u>A</u> ACGCTGGGTACGG						
ApjSNP059_CT	WD repeat and FYVE domain-containing protein 3	F:TTCCAGGGATTTGACAGAGG	110	0.530	0.500	С	0.446	0.984
		R:TGGCATCTAAAGCTGCTAGTCT						
		P:TCCAGGAGAGATCCTAGGGTGTACTGGG						
ApjSNP060_AT	similar to LOC398543 protein	F:CCACTACACATCGGTGACCA	110	0.095	0.433	А	0.309	0.008 *
		R:CATCTCCTTCCGATAACACAGTT						
		P:AGATGAAGAATGTATTAACGCTGCACACT						
ApjSNP061_CT	Coiled-coil domain-containing protein C6orf97	F:GCTGTTGCCGATGAAACAAT	110	0.479	0.447	Т	0.329	0.489
		R:CAAATTGAACGAGATGGAGACA						
		P:AGAATATCCTGCCT <u>T</u> GGGATAACGTAAACC						
ApjSNP062_CT	Uncharacterized gene 48 protein	F:CAGAAGGATAAAGTCCAAGAGACC	86	0.182	0.220	Т	0.198	0.809
		R:TTCTCCTTTCTGTCCATCCTG						
		P:ACAGGCCTATAGCTACGATCAGGAATCG						
ApjSNP063_AT	Uncharacterized protein C2orf73 homolog	F:CACATGTGTCACCTCTGGCTA	73	0.479	0.586	А	0.311	0.252
		R:ACTGGAACAGCGCCTTTAGA						
		P:CAGCTCAAACCCTCACAACTATGCAAG						

Table 2. Cont.

Locus ID	Gene Name	Primers and Probes (5'-3')	Size (bp)	Ho	He	MA	MAF	<i>p</i> -Value
ApjSNP064_AT	Methionine synthase	F:TCGATACCCTTCACCAAAGAAT	103	0.486	0.495	Т	0.432	0.654
		R:CGAGGGTCTTGGGAAAGGA						
		P:CCAGGCTTCATCATCAACAGCTTTCTAA						
ApjSNP065_CT	Tubulin alpha chain	F:CATAGCTTCGGTGGTGGAAC	85	0.061	0.091	Т	0.091	1.000
		R:GCTTCGATTTCTTGCCGTAG						
		P:GGATTTGCAG <u>C</u> TCTACTTCTTGAACGCG						
ApjSNP066_AG	TATA element modulatory factor	F:TGGTGCTCAGCTGAATCTGT	86	0.415	0.100	G	0.321	0.007 *
		R:TGGTCTCTTCGTGAGCCTCT						
		P:GAAACAACAAGACAAGACCTCGAGAGGCTT						
ApjSNP067_AG	TATA element modulatory factor	F:GCAACTGGAGGCAGAGAGAG	79	0.315	0.400	А	0.206	0.486
		R:GGCCTGCTCGAGTTTACCT						
		P:AGAGACCAAGGA <u>A</u> GAGCTGGAAGAGAA						
ApjSNP068_CT	Uncharacterized protein KIAA1704 homolog	F:TGACACCTATGGACCGTCTCT	90	0.412	0.504	С	0.477	0.135
		R:GGAGGTAATGGTGGACCAAA						
		P:GGATTCAAAGGTGT <u>C</u> GACAAAGAGTCTGAAC						
ApjSNP069_CT	WD repeat-containing protein KIAA1875	F:GGGTCTTCCAGCCAATGATA	103	0.326	0.225	С	0.471	0.390
		R:ACCACGGCTACGTTTGAGTC						
		P:TACTGGTTGATCGCTCTGGAAGAAACAGGA						
ApjSNP070_AG	Glycoprotein 3-alpha-L-fucosyltransferase A	F:CCAGGAAGGGGTAGACTTGC	102	0.528	0.469	G	0.334	0.051
		R:ATCTCGCCGTTCAAGTTGTT						
		P:CTCAGGAAGTTCT <u>A</u> GAGAGGAAGGATGTC						
ApjSNP071_AG	N/A	F:CGAAACTATAGTGACCTCTTGGTTA	104	0.364	0.470	А	0.364	0.189
		R:CAAGCCCTAGTCTCTTCATTCG						
		P:CAGAATTTCTCTC <u>G</u> AAGTCCTTTGCCAG						
ApjSNP072_AG	N/A	F:GAGTTAGACCCTCGGCTAGGTA	87	0.388	0.412	G	0.333	0.247
		R:GCAAAGAGCCTAGCCTTTAGGT						
		P:TGCATCAGTACTAGCAGCATGGAAAACT						

Table 2. Cont.

Locus ID	Gene name	Primers and probes (5'-3')	Size (bp)	Ho	He	MA	MAF	<i>p</i> -Value
ApjSNP073_AG	N/A	F:AAATGTACAGACCCGCATGA	107	0.225	0.309	А	0.188	0.104
		R:CTGGAAAAACAGTGTGAACCAA						
		P:TGTAAAATTAATGA <u>G</u> CCGTTCGAACCAAGAG						
ApjSNP074_AT	N/A	F:GATGGTGAAAATCACGGAGAA	103	0.300	0.404	А	0.275	0.108
		R:TTCTATGTCTTGTTGATGCAGAGAC						
		P:CACAATAACCTGGAAATATCAACCTTAGAAGAATTCA						
ApjSNP075_AT	N/A	F:GACCACGATGACAGCCAGTA	95	0.630	0.879	Т	0.450	0.328
		R:CTCGCCAAGTCAGGAAAAAG						
		P:AGGATCGTCATTCGGGCACTCTTGG						
ApjSNP076_CT	N/A	F:AACTCTCGATGGAATGCAAAG	108	0.175	0.392	С	0.263	0.001 *
		R:AACAGACTCGGTCGCATCTC						
		P:GATAGTTCTGACAGCGATTTAGGAGACTAA						
ApjSNP077_CT	N/A	F:AACCATCCTGTAGCGAAACC	103	0.175	0.339	Т	0.213	0.004 *
		R:CGGGGACGAGGATATTGTTA						
		P:GTGTTGAATGAAG <u>T</u> CGTTCGCGTAAATGC						
ApjSNP078_GT	N/A	F:GCCAAGCAACATACAGAAGGA	87	0.563	0.907	Т	0.487	0.637
		R:TAGTTGGGCTGTCTTGCTGA						
		P:TTGCTGCATTAATGT <u>T</u> TAGATGATGATGTGTCT						
ApjSNP079_AG	N/A	F:TGGGCAGAAGAAAATTTGGA	99	0.475	0.469	G	0.375	0.084
		R:GAGTGGCACATGACTTGGTG						
		P:CTGCAATTGGACAACCCCATGCTCAT						
ApjSNP080_CT	N/A	F:GGGCGCTATCAGACTTTGAC	110	0.200	0.292	С	0.175	0.062
		R:GCACCCTCTATTTTAGCTGTTCA						
		P:TCTTGCTAGCTAA <u>T</u> GGGAAAGAACGTTAT						
ApjSNP081_CT	N/A	F:CTGGTTGCAATAGGTTATTTGG	103	0.075	0.073	С	0.038	0.780
		R:TGAATACATGCCGTTTCTGA						
		P:GTTGGATTCAGAACAGACTGCCATTCC						

Table 2. Cont.

Locus ID	Gene Name	Primers and Probes (5'–3')	Size (bp)	Ho	He	MA	MAF	<i>p</i> -Value
ApjSNP082_CT	N/A	F:CAGAAACGGCATGTATTCAAAC	94	0.500	0.404	С	0.275	0.098
		R:CCCGACCACAAGGAAAGATA						
		P:AGGGGAGTTTGTGA <u>T</u> GACAAATTGTTGCAG						
ApjSNP083_AC	N/A	F:CACGATGCCCTGTGTGTAAT	108	0.325	0.453	С	0.338	0.073
		R:GTCGGCCTCCTGACTAACAG						
		P:GCGCAGCAG <u>A</u> AACGGCGTGGA						
ApjSNP084_CG	N/A	F:GGGTGGTGCATTTTCTTCAT	75	0.150	0.444	G	0.325	0.000 *
		R:TGGCTTCAGTTACACCATCCT						
		P:ATCCTTGTGGTC <u>G</u> CCTGATCTTGTGTT						
ApjSNP085_AG	N/A	F:CGTCATTCGCTCCAAATACC	110	0.833	0.896	А	0.487	0.093
		R:GTCGTAGAGAGACATAACGATAACTGA						
		P:CCATAATGCATA <u>G</u> TGGCTGCAGCATAA						
ApjSNP086_AG	N/A	F:CGACAATATACTACAAATGCCCTGT	83	0.050	0.461	G	0.350	0.000 *
		R:GATGATGAATGGGTTGTTTGTG						
		P:CAAGGCGAGTTC <u>G</u> TCACACGAAAAGT						
ApjSNP087_AC	N/A	F:CACTCTGGCCTTGCACTCTT	109	0.450	0.353	С	0.225	0.252
		R:TGTGAGAACAATAGGTTCACAGGT						
		P:GGGCAAACTGATGT <u>C</u> ATGTTCACAGGTATGT						
ApjSNP088_AG	N/A	F:ATGAAGCATGCGTGAATGAG	83	0.250	0.222	G	0.125	0.256
		R:CGATTTCACTGCTGTCATCAA						
		P:AACTGTGGAGATGGTA <u>A</u> CATATTCTATGAAGAGAA						
ApjSNP089_AG	N/A	F:TGGTGAGAAGCATCCACAGA	93	0.325	0.468	А	0.363	0.051
		R:GTTGTTTTGAAGGCACTGATGA						
		P:AAGTTCTTAAATGC <u>A</u> GAACTGGGTCAGAACA						
ApjSNP090_CT	N/A	F:TTGTACCGAGAAAGGGATGTTT	110	0.161	0.373	Т	0.242	0.002 *
		R:CCTGAACAACATCTGCCTGA						
		P:AGAGTATATTTCAAA <u>C</u> GAAAACGGGAGTAGGGT						

Table 2. Cont.

Locus ID	Gene Name	Primers and Probes (5'–3')	Size (bp)	Но	He	MA	MAF	<i>p</i> -Value
ApjSNP091_CT	N/A	F:TGCGTCATTCTAACCAACCA	102	0.200	0.380	С	0.250	0.004 *
		R :AACACTTATGTAGGCGAGTCTTGA						
		P:CAAAGCGCTTCA <u>T</u> TTTCACAGCAACTA						
ApjSNP092_CT	N/A	F:TGACTGGACGTCAGATGTGG	81	0.075	0.073	С	0.038	0.780
		R:GTGGGCTTCCAGACACAGAT						
		P:GGTTGCATCAAGGTCCCTGGGTACATACA						
ApjSNP093_AG	N/A	F:TGAAATGTGGTGTGACTTGC	80	0.222	0.282	G	0.167	0.227
		R:TGTGTGACTTCAGCATCTCTGT						
		P:GAATTGTATAATTGG <u>A</u> TGCTGTGTGTCACTTAT						
ApjSNP094_GT	N/A	F:TCTGCTAAGTTGTTGAGAGGATG	108	0.171	0.358	Т	0.229	0.003 *
		R:CGAACGGTTGGTATTTGTGA						
		P:TTCTGGTCACTTGCCCCAGGTTCCAC						
ApjSNP095_AG	N/A	F:ATTTGCGGCTCTTCTGTTCA	110	0.225	0.367	А	0.238	0.018
		R :TGAAGTGAACTCACCCACGA						
		P:AAACTTGGCAACG <u>A</u> AGACGTCAGCAT						
ApjSNP096_CT	N/A	F:TCATTCCTGTATTGCTACTACTCTGTG	109	0.333	0.491	С	0.409	0.060
		R :TGTGGTATGCCCATCGATTT						
		P:TAAACAATAGTACTTAA <u>T</u> GGCATTGAAGACAACAAAC						
ApjSNP097_CG	N/A	F:CACAGTGATGTGTATGTACGTTCG	94	0.316	0.337	С	0.211	0.698
		R:GACCTTCGCTTTGTGCCTAC						
		P:ACACACCGTATATACCGAAATCTGGAAATTATCTT						
ApjSNP098_CT	N/A	F:CTGTGTCAGAGAGGAAGAGTGC	107	0.158	0.147	С	0.079	0.512
		R:CGAAAGCTATTTCAAACCCAGT						
		P:GGGTACTATCAAAA <u>T</u> TGACTCACAAAGCGAC						
ApjSNP099_AG	N/A	F:GACCTTCTGCTCTGCCTGAC	97	0.075	0.162	G	0.088	0.080
		R:CGGATATCAACAAACCAGAGC						
		P:TCCTCATCTTC <u>G</u> GTGTCTTGCGAAC						

Table 2. Cont.

Locus ID	Gene Name	Primers and Probes (5'-3')	Size (bp)	Ho	He	MA	MAF	<i>p</i> -Value
ApjSNP100_GT	N/A	F:TCCACTGAGCCATCCTGATT	103	0.505	0.547	Т	0.458	0.279
		R:GAAGAAAAACATGTCCCGATG						
		P:AGTGGCTCCCCCTGGAATGTAATCCTG						
ApjSNP101_GT	N/A	F:CTGCTGAAGTATGACAACATTAGAGAC	109	0.075	0.240	Т	0.138	0.000 *
		R:CTAGTACTTTCTTCTTCAGTAGTTGG						
		P:CTATTGAAAGCTCGATAGGCACATCCTG						

The underlined bases in the probe sequences indicated the positions of the SNPs; H_0 , observed heterozygosity; H_e , expected heterozygosity; MA, minor allele; MAF, minor allele frequency;

*P*_{HWE}, *P* values for Hardy–Weinberg equilibrium (HWE) test; *, statistically significant after sequential Bonferroni correction.

As the gene-derived SNPs reside in or are immediately next to protein-coding sequences, they stand a better chance for identifying functional genes that are responsible for complex traits as well as simply inherited traits [8,9]. In our study, 70 SNP markers (Table 2) were developed from the EST sequences showing significant similarity to an entry in the NCBI nr database [10]. Among the annotation information, genes potentially involved in growth or immunity (e.g., epidermal growth factor receptor, Zinc finger protein 62 homolog and heat shock protein 90 kDa beta) were identified. It would be interesting to see whether any of these growth- or immune-related SNPs are highlighted in future QTL mapping of economically important traits, such as high growth rate and disease resistance.

3. Experimental Section

3.1. Sampling and DNA Extraction

A total of 48 *A. japonicus* individuals used for SNP marker validation were collected from four natural populations (Dalian, Yantai, Qingdao and Wendeng) in China. Genetic parameters of the validated SNP markers were further evaluated using 40 *A. japonicus* individuals from the Rongcheng (Shandong, China) population. Genomic DNA was extracted from the muscles of sea cucumbers by following the protocol developed by Zhan *et al.* [11]. The quantity and integrity of genomic DNA was determined using an UltrospecTM 2100 pro UV/Visible Spectrophotometer (Amersham Biosciences, Uppsala, Sweden) and gel electrophoresis, respectively.

3.2. SNP Discovery and Genotyping

Our group has recently released a large amount of transcriptomic data by 454 sequencing of eight cDNA libraries constructed using more than 200 sea cucumber individuals. Potential SNPs were detected from the assembled contigs using the program GS Reference Mapper (version 2.6, Roche 454 Life Sciences: Branford, CT, USA, 2011) with default parameters (cDNA mode). More than 54,000 putative SNPs were identified from the dataset, 200 of which were selected in this study for marker development with the selection criteria of at least $3 \times$ occurrence of the minority allele and at least $6 \times$ contigs coverage (number of reads forming the contig). SNP genotyping was performed using a recently developed cost-effective HRM method [12]. For each locus, three non-modified oligonucleotides were used, corresponding to two PCR primers and one probe, primers were designed using Primer3 [13] with the following rules: (1) primer length should be at least 20 bases; (2) product size should not exceed 120 bp in order to decrease the probability of intron interference; (3) the primer Tm should be between 59 °C and 61 °C; (4) the primer GC% should be 40%-60%; and (5) the amplicon contains only one SNP site. Probes were designed using OligoCalc [14] with the following criteria: (1) SNP site locates in the middle of the probe; (2) the length of probe is between 20 and 35 bases; (3) Tm is about 60 °C; (4) the 3' end of each probe is blocked by two mismatch bases; and (5) no overlap between primes and probe. Each SNP locus was first amplified by an asymmetrical PCR with HRM fluorescent dye in the PCR master mix and then interrogated by an unlabeled probe. The 48 individuals of A. japonicus collected from four natural populations were used for SNP marker validation. PCR amplifications were carried out in a 10 µL reaction mixture containing 20 ng of genomic DNA, 1× PCR buffer, 0.2 mM dNTPs, 1.5 mM MgCl₂, 0.5 U Taq DNA polymerase (Takara, Dalian, China), 0.1 μ M forward primer, 0.5 μ M reverse primer and 1× LCGreen Plus (Idaho technology inc., Salt Lake City, Utah, USA). The amplification was programmed as: an initial denaturation at 95 °C for 5 min, followed by 55 cycles of 95 °C for 40 s, 60 °C for 40 s and 72 °C for 40 s, finishing with a final elongation at 72 °C for 5 min. The PCR products were checked by gel electrophoresis, and those with correct PCR product sizes were then subjected to probe testing. An aliquot of the appropriate probe was added in each reaction to a final concentration of 5 μ M. The PCR product and probe mixture were denatured at 95 °C for 15 min and then slowly cooled to 4 °C. HRM genotyping was immediately performed on a Light Scanner instrument (HR96 model, Idaho technology inc., Salt Lake City, Utah, USA) with continuous melting curve acquisition (10 acquisitions per °C) during a 0.1 °C/s ramp from 40 to 95 °C.

3.3. Data Analysis

Data were retrieved and analyzed using the Light Scanner software followed by manual curation of the obtained genotype calls. POPGENE [15] was used to analyze allele frequency, expected (*He*) and observed (*Ho*) heterozygosities, and tests for deviation from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD).

4. Conclusions

In summary, 101 gene-based SNPs were successfully developed from the transcriptome sequences of *A. japonicus*. These developed markers are expected to be useful for future QTL analysis, and to facilitate MAS in *A. japonicus*.

Acknowledgments

Financial support for this work was provided by the National Key Technology R&D Program of China (2011BAD13B05 and 2011BAD13B06), and the National High Technology Research and Development Program of China (2012AA10A412).

References

- 1. Sloan, N.A. References. In *Echinodermata*; Keegan, B.F., O'Connor, B.D.S., Eds.; A.A. Balkema: Rotterdam, The Netherlands, 1985; pp. 109–124.
- Chen, J.X. Present Status and Prospects of Sea Cucumber Industry in China. In Advances in Sea Cucumber Aquaculture and Management; Lovatelli, A., Conand, C., Purcell, S., Uthicke, S., Hamel, J.F., Mercier, A., Eds.; FAO: Rome, Italy, 2004; pp. 25–38.
- 3. Collard, B.C.Y.; Jahufer, M.Z.Z.; Brouwer, J.B.; Pang, E.C.K. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica* **2005**, *142*, 169–196.
- 4. Cho, R.J.; Mindrinos, M.; Richards, D.R.; Sapolsky, R.J.; Anderson, M.; Drenkard, E.; Dewdney, J.; Reuber, T.L.; Stammers, M.; Federspiel, N.; *et al.* Genome-wide mapping with biallelic markers in *Arabidopsis thaliana*. *Nat. Genet.* **1999**, *23*, 203–207.

- 5. Sun, W.J.; Li, Q.; Kong, L.F. Characterization of thirteen single nucleotide polymorphism markers in the sea cucumber (*Apostichopus japonicus*). *Conserv. Genet. Resour.* **2010**, *2*, 141–144.
- 6. Yang, A.F.; Sun, D.F.; Liu S.K.; Dong, Y., Chen, Z.; Zhou, Z.C. Characterization of fifteen SNP markers by mining EST in sea cucumber, *Apostichopus japonicus*. J. Genet. **2012**, *91*, e49–e53.
- Du, H.X.; Bao, Z.M.; Hou, R.; Wang, S.; Su, H.L.; Yan, J.J.; Tian, M.L.; Li, Y.; Wei, W.; Hu, X.L.; *et al.* Transcriptome sequencing and characterization for the sea cucumber *Apostichopus japonicus* (Selenka, 1867). *PLoS One* 2012, 7, doi:10.1371/journal.pone.0033311.
- Tsumura, Y.; Kado, T.; Takahashi, T.; Tani, N.; Ujino-Ihara, T.; Iwata, H. Genome scan to detect genetic structure and adaptive genes of natural populations of *Cryptomeria japonica*. *Genetics* 2007, 176, 2393–2403.
- Namroud, M.C.; Beaulieu, J.; Juge, N.; Laroche, J.; Bousquet, J. Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in white spruce. *Mol. Ecol.* 2008, 17, 3599–3613.
- 10. NCBI nr database. Available online: ftp://ftp.ncbi.nlm.nih.gov/blast/db/ (accessed on 25 January 2012).
- Zhan, A.B.; Bao, Z.M.; Lu, W.; Hu, X.L.; Peng, W.; Wang, M.L.; Hu, J.J. Development and characterization of 45 novel microsatellite markers for sea cucumber (*Apostichopus japonicus*). *Mol. Ecol. Notes* 2007, 7, 1345–1348.
- Wang, S.; Zhang, L.L.; Meyer, E.; Matz, M.V. Construction of a high-resolution genetic linkage map and comparative genome analysis for the reef-building coral *Acropora millepora*. *Genome Biol.* 2009, 10, doi:10.1186/gb-2009-10-11-r126.
- 13. Rozen, S.; Skaletsky, H. References. In *Bioinformatics Methods and Protocols: Methods in Molecular Biology*; Krawetz, S., Misener, S., Eds.; Humana: Totowa, NJ, USA, 2000; pp. 365–386.
- 14. Kibbe, W.A. OligoCalc: An online oligonucleotide properties calculator. *Nucleic Acids Res.* 2007, *35*, 43–46.
- 15. Yeh, F.C.; Boyle, T.J.B. Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belg. J. Bot.* **1997**, *129*, 157.

© 2012 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).