

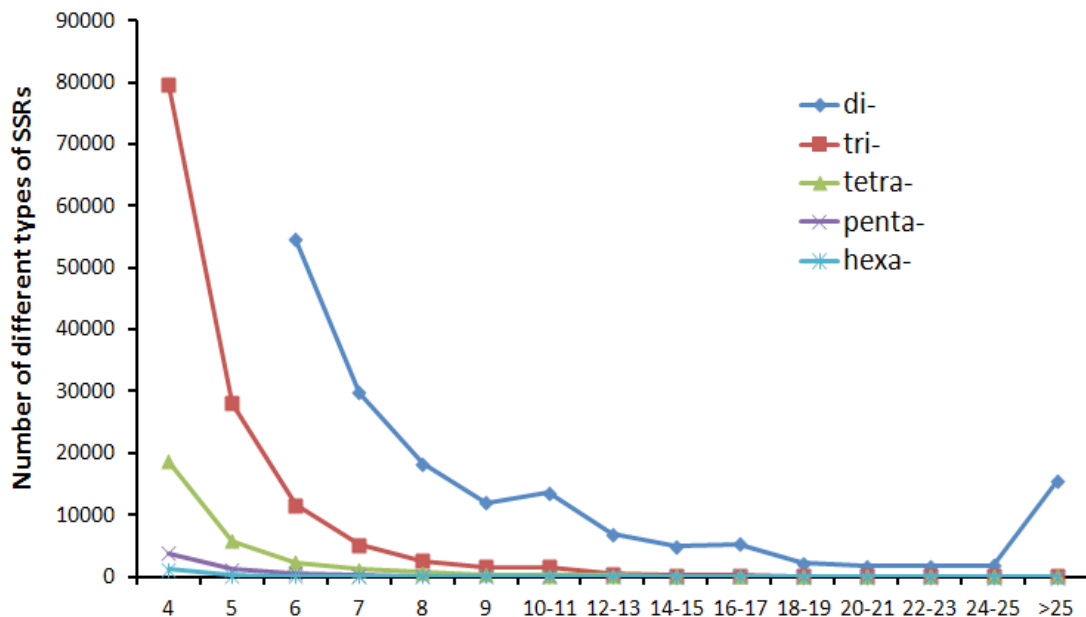
Supplementary Material

Genome Survey and Large-scale Isolation of Microsatellites through Illumina Sequencing from the Red Swamp Crayfish *Procambarus clarkii*

Xiao Zhang¹, Guoqing Shen¹, Yang Wang¹, Pengdan Huang^{1,2}, Yan Shui², Zenghong Xu² and Huaishun Shen^{1,2*}

¹Wuxi Fisheries College, Nanjing Agricultural University, Nanjing, 210095, PR China

²Key Laboratory of Freshwater Fisheries and Germplasm Resources Utilization, Ministry of Agriculture, Freshwater Fisheries Research Center, Chinese Academy of Fishery Sciences, No. 9 Shanshui East Road, Wuxi 214081, PR China



Supplementary Fig. S1. Observed number of microsatellites with di-, tri-, tetra-, penta- and hexa-nucleotide motifs in 1,625Mb sequence.

* Corresponding author: shenhuaishun@ffrc.cn
0030-9923/2023/0002-795 \$ 9.00/0



Copyright 2023 by the authors. Licensee Zoological Society of Pakistan. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Supplementary Table 1. Sequencing data output statistics.

Library	Insert size (bp)	Raw reads (bp)	Raw Base (bp)	Effective rate (%)	Clean base (bp)	Error rate (%)	Q20 (%)	Q30 (%)	GC content (%)
DES02111-S	350bp	401,986,051	120,595,815,300	100	120,595,815,300	0.05	93.78	86.75	42.86

Supplementary Table 2. Statistics of shrimp genome characteristics obtained by Kmer=17 analysis.

K-mer	K-mer number	K-mer depth	Genome size (Mbp)	Revised genome size (Mbp)	Heterozygous ratio (%)	Repeat (%)
17	83,427,961,764	23	3627.30	3579.38	0.88	77.65

K-mer, analyzes the size of the selected kmer; K-mer number, the total number of kmers obtained by Soapdenovo software; K-mer depth, kmer depth, which is the expected value corresponding to the Poisson distribution; Genome size, calculated genome size, K-mer number/K-mer Depth, in M (mega); Revised genome size, the corrected genome size, eliminating the error caused by the error kmer, in M (mega); Heterozygous ratio, the heterozygosity ratio, calculated according to the formula (7) of 5.1; Repeat: The repetition rate obtained by kmer analysis is calculated as a percentage of the total number of kmer of 1.8 times the main peak as a percentage of all kmer numbers.

Supplementary Table 3. Contig assembly result statistics.

Total length (bp)	Total number	Max length (bp)	N50 length (bp)	N90 length (bp)
1,535,295,243	3,711,514	61,160	879	130

Supplementary Table 4. Scaffold assembly result statistics.

Total length (bp)	Total number	Max length (bp)	N50 length (bp)	N90 length (bp)
1,682,115,001	2,940,268	81,806	1,426	159

Supplementary Table 5. The genetic parameters of 10 individuals of *Procambarus clarkii* in 22 microsatellite loci.

No. Loci	Primer1	tm of primer 1	Primer 2	tm of primer 2	Repeat unit of SSR	num-ber of alleles	Ho	He	PIC
1	scaffold137841 11.7	AGCAGTGCATCACCATACCAAITCTG	GGACAGTCTTCTTCTGCTGTCTCC	63.02	(AGC)10	3	0.5	0.43	0.37
2	scaffold1524733 14.4	CAACGGGAGATTGGGTATGTCTAA	TATATGAGACGGTTTACACGGGGCT	62.95	(TC)12	5	1	0.8	0.7
3	scaffold1526357 14.2	GCTTCATGTTGGGATTTCTTGGATA	CTCTGGGGTGGCTTAATCTCCAT	62.79	(GATT)11	5	0.7	0.75	0.67
4	scaffold1893661 16.2	TGTGTGCTGATTCATGTGTGTACC	CAGAATGTTTGTACTGTAAACGGCTGC	62.2	(AAT)10	7	1	0.77	0.69
5	scaffold2027925 11.2	GTGCCCCAACCAACCACCTATATATC	TGTAGTATCTGTGCTTACGACGGCC	62.69	(ACC)10	4	0.5	0.75	0.66
6	scaffold2039339 11.1	TCTTGTGTTGTACCTCAGTCTGGTCC	AGCTCATAGGTTCCGAATCCGTATC	62.93	(ATT)10	2	0.7	0.48	0.35
7	scaffold2055415 12.4	GAAAGACAGGGAAATGTGTGTGTTG	TTTCTTCCATCCTTCTTCCCTTCC	63.05	(GAA)10	4	0.2	0.49	0.42
8	scaffold256173 15.6	TGACTGTCTTGCAGGCAGAACTAT	TCTGTAGTATTACACCGCGCACAT	62.9	(TACAC)10	4	0.7	0.71	0.61
9	scaffold260377 14.0	CACGAGAAAATTTGTCAITCATTACGTC	TGATATGATAGTTGGGTTGGGTT	62.59	(AACT)10	3	0.5	0.51	0.41
10	scaffold1022879 17.4	CAATGTTTTTTTAAGCGTTGTTGC	TTGGATCAGTGTTTTTGCCAAGTGT	62.96	(TGTA)10	3	0.6	0.7	0.59
11	scaffold1188287 13.9	CATTTTGACATGCAAACAAATTTGA	TACGCTAGACTTGTTAGCGGCTTT	62.7	(ATTG)10	6	0.6	0.82	0.75
12	scaffold1351069 13.5	GAGGGAAGGTTGAAGAGAAAAGAG	CCTACTATGGCCTTGTCCATCTG	63.12	(GGA)10	4	0.8	0.77	0.61
13	scaffold1598717 14.0	CCGTGTATAATTTGGCTTAGGTTAAGG	TGAAAAGAAGCTGCAAAITTACTGCCA	63.07	(TAGGT)10	4	0.7	0.6	0.51
14	scaffold1867767 15.6	GAAATCAAAAACCAAAACAGGGTTGTC	GTTTAGGCTCAGTTGAGGCTTGC	63.68	(AGACAC)7	4	0.4	0.67	0.57
15	scaffold2029901 9.5	CCCAGTGTGGAGACTGTGTCTT	AGTTGGTTTAGTGCATACAGAGTCCA	62.37	(AACCT)10	2	0.4	0.34	0.27
16	scaffold4888809 14.2	GGTATTTATCCTCGTGTGGCTTACG	CCCCCTATTCGCTACTGACTTAT	62.06	(GTATAA)10	4	0.8	0.7	0.6
17	scaffold1940553 16.4	CGATGTTTCGCCCTCCAAITCTAATA	CCAGAGGATGTCTTCTTGGACAGT	63.11	(AGTAC)9	3	0.5	0.54	0.46
18	scaffold1092139 15.3	TAAACGATTTCCCAAAAAATCAGGTT	CATCAATAATCCCATTTGTACCCTTCTCA	62.33	(TTTA)11	3	0.4	0.68	0.57
19	scaffold1946501 10.0	GAGAGAAAGGAAAAGGAGGACACAG	TCACACAATCTGTTTTCCCTTGGAG	62.59	(AGACAG)9	2	0.3	0.52	0.37
20	scaffold796857 10.6	GGGAATGTTGCACACTGTGTTGG	CTCTTTCCAAAGCTTCCACTCCCTCT	63.642	(GGGAGA)11	3	0.3	0.56	0.44
21	scaffold1792375 12.7	TTAACACAGAGCCCTCTTTTTGGCT	GCCAGAGAAITTAGGACCAAAATCAA	62.754	(TAGGT)12	3	0.5	0.51	0.41
22	scaffold770215 12.6	AGCAGCTGATGTCTCAAGGAAAAGT	GACGAATTTGTTGGCCATGTG	62.939	(GCA)10	2	0.2	0.44	0.33
Mean						3.6364	0.6	0.62	0.52

Ho, observed heterozygosity; He, expected heterozygosity; PIC, polymorphism information content.