Optimization of high molecular weight DNA extraction methods in shrimp for a long-read sequencing platform

Pacharaporn Angthong[1], Tanaporn Uengwetwanit[1], Wirulda Pootakham[2], Kanchana Sittikankaew[1], Chutima Sonthirod[2], Duangjai Sangsrakru[2], Thippawan Yoocha[2], Intawat Nookaew[3], Thidathip Wongsurawat[4], Piroon Jenjaroenpun[4], Wanilada Rungrassamee[1] and Nitsara Karoonuthaisiri[1]

[1] Microarray Research Team, National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency, Pathum Thani, Thailand
[2] National Omics Center, National Science and Technology Development Agency, Pathum Thani, Thailand
[3] Department of Biomedical Informatics, College of Medicine, University of Arkansas for Medical Sciences, Little Rock, AR, United States of America
[4] Division of Bioinformatics and Data Management for Research, Department of Research and Development, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand

ABSTRACT

Marine organisms are important to global food security as they are the largest source of animal proteins feeding mankind. Genomics-assisted aquaculture can increase yield while preserving the environment to ensure sufficient and sustainable production for global food security. However, only few high-quality genome sequences of marine organisms, especially shellfish, are available to the public partly because of the difficulty in the sequence assembly due to the complex nature of their genomes. A key step for a successful genome sequencing is the preparation of high-quality high molecular weight (HMW) genomic DNA. This study evaluated the effectiveness of five DNA extraction protocols (CTAB, Genomic-tip, Mollusc DNA, TIANamp Marine Animals DNA, and Sbeadex livestock kits) in obtaining shrimp HMW DNA for a long-read sequencing platform. DNA samples were assessed for quality and quantity using a Qubit fluorometer, NanoDrop spectrophotometer and pulsed-field gel electrophoresis. Among the five extraction methods examined without further optimization, the Genomic-tip kit yielded genomic DNA with the highest quality. However, further modifications of these established protocols might yield even better DNA quality and quantity. To further investigate whether the obtained genomic DNA could be used in a long-read sequencing application, DNA samples from the top three extraction methods (CTAB method, Genomic-tip and Mollusc DNA kits) were used for Pacific Biosciences (PacBio) library construction and sequencing. Genomic DNA obtained from Genomic-tip and Mollusc DNA kits allowed successful library construction, while the DNA obtained from the CTAB method did not. Genomic DNA isolated using the Genomic-tip kit yielded a higher number of long reads (N50 of 14.57 Kb) than those obtained from Mollusc DNA kits (N50 of 9.74 Kb). Thus, this study identified an effective extraction method for high-quality HMW genomic DNA of shrimp that can be applied to other marine organisms for a long-read sequencing platform.

Corresponding author
Nitsara Karoonuthaisiri,
nitsara@alumni.stanford.edu,
nitsara.kar@biotec.or.th

INTRODUCTION

Seafood is the largest source of animal protein (*FAO, 2016*), making aquaculture the world's fastest growing food-production (*Anderson et al., 2017*). The industry strives to increase production capacity while reducing cost by utilizing modern technologies. Genome sequence can serve as an important tool for the identification of key pathways to guide potential functions (*Collins et al., 2003*; *Ramanan et al., 2012*). Genome sequence has enabled the identification of single nucleotide polymorphisms (SNPs) and genetic markers for desirable traits (*Zwane et al., 2019*) such as productivity, disease resistance and reproductive maturation (*Guppy et al., 2018*). For instance, specific SNPs for growth performance and immunity traits have been identified and utilized for breeding programs assisted by genomic selection in Blanco Orejinegro cattle (*Martínez et al., 2010*; *Martínez et al., 2016*), chicken (*Li et al., 2013*; *Xie et al., 2012*) and pig (*Geraci et al., 2019*; *Yang et al., 2019*). Genomics has indeed contributed to the exponential growth of terrestrial animal production, yet it has not been fully exploited in aquaculture industry due to several challenges in obtaining whole genome sequences of marine organisms.

With continuing progress of sequencing technology, whole genome assemblies of few marine organisms are now available such as Atlantic salmon (*Salmo salar*) (*Davidson et al., 2010*), Atlantic cod (*Gadus morhua*) (*Star et al., 2011*), Pacific oyster (*Crassostrea gigas*) (*Zhang et al., 2012*), Pacific white shrimp (*Litopenaeus vannamei*) (*Zhang et al., 2019*), Black tiger shrimp (*Penaeus monodon*) (*Uengwetwanit et al., 2020*) and Rainbow trout (*Oncorhynchus mykiss*) (*Berthelot et al., 2014*). Although whole genome sequencing and de novo assembly have become a routine method for functional genomics studies, genome sequencing of non-model organisms with a large proportion of long repetitive sequences such as marine invertebrates remains challenging. Recently, the long-read sequencing platforms present an alternative approach to overcome those challenges with the following advantages: (i) longer read length to facilitate de novo assembly and enable direct detection of haplotypes and even whole chromosome phasing; (ii) higher consensus accuracy to enable rare variant detection; (iii) low cost per run; and (iv) provide the epigenetics information (*Metzker, 2010*; *Schadt, Turner & Kasarskis, 2010*).

With an ability to span repetitive sequences, long-read sequencing technology presents a promising alternative to obtain genome assemblies for marine species (*Amarasinghe et al., 2020*; *Pollard et al., 2018*). High-quality genomic DNA is the key to allow the utilization of the long-read sequencing technology. However, obtaining such high-quality genomic DNA from crustaceans faces several challenges because crustacean DNA is easily degraded due to endonuclease enzyme (*Rahman, Schmidt & Hughes, 2017*) and the purity can be jeopardized by high amounts of polysaccharides and polyphenolic proteins, which can inhibit subsequent molecular applications (*Panova et al., 2016*).

Despite many DNA extraction methods available, the efficiency of methods depends on specimen and preservation (*Abdel-Latif & Osman, 2017*; *Psifidi et al., 2015*). DNA extraction methods in marine organisms have been evaluated for PCR application (*Chakraborty, Saha & Ananthram, 2020*; *Jasrotia & Langer, 2019*; *Nadia & Sadok, 2017*; *Pereira et al., 2011*) but not for long-read sequencing.

This study therefore aimed to evaluate DNA extraction methods that yielded high quality genomic DNA, enabling full exploitation of the long-read sequencing technology. We employed the black tiger shrimp (*Penaeus monodon*) as a model marine organism as it is one of the most widely cultured shrimp species in the world, accounting for 52% of total global farm production (*FAO, 2016*).

MATERIALS & METHODS

Shrimp collection

Four-month-old black tiger shrimps ($n = 4$) were obtained from a local farm in Prachuap Khiri Khan, Thailand. Muscle tissue from each shrimp was collected and immediately frozen in liquid nitrogen and stored at $-80\,°C$ prior to DNA extraction.

Genomic DNA extraction methods

Five DNA extraction methods namely (I) CTAB method, (II) Genomic-tip 100/G kit (Qiagen, Germany), (III) E.Z.N.A.® Mollusc DNA kit (Omega bio-tek, USA), (IV) TIANamp Marine Animals DNA kit (Tinagen, China) and (V) Sbeadex livestock kit (LGC, Germany) were evaluated for genomic DNA preparation (4 biological replicates/extraction method).

Frozen shrimp muscle was transferred to mortar containing liquid nitrogen and pulverized. All tissue samples were immediately processed according to the instruction of each method as follows:

Cetyltrimethyl ammonium bromide (CTAB) method

CTAB method used in this study followed a previously described protocol (*Winnepenninckx, Backeljau & De Wachte, 1993*). Dried powder of frozen shrimp tissue (100.02 ± 0.04 mg) was mixed with 800 μL of an extraction buffer (2% (w/v) CTAB, 1% (w/v) polyvinylpyrpolidone, 1.4 M NaCl, 20 mM EDTA, pH 8.0; 100 mM Tris–HCl, pH 8.0) by gently inverting the tube before 20% SDS (20 μL) was added and incubated at $65\,°C$ for 10 min. RNA was removed by incubating with 10 μL of 10 mg/mL RNase A (final concentration of 0.12 mg/mL) at $37\,°C$ for 30 min. Subsequently, 50 μL of β-mercaptoethanol and 500 μL of chloroform were added and mixed by inverting. The mixture was centrifuged at 10,000 g for 5 min at $10\,°C$. The upper aqueous phase was transferred and extracted with 500 μL of chloroform:isoamyl alcohol (24:1) by inverting and centrifugation as before. The chloroform:isoamyl alcohol extraction step was repeated twice before DNA was precipitated with 2 volumes of 100% ethanol. The sample was centrifuged at 12,000 g for 10 min at $10\,°C$ and the supernatant was discarded. The DNA pellet was washed with one mL of 70% ethanol, centrifuged at 12,000 g for 5 min at $10\,°C$, air-dried and resuspended in 50 μL of TE buffer (10 mM Tris–HCl, pH 8.0; 1 mM EDTA, pH 8.0).

Genomic-tip kit

DNA was extracted according to the manufacturer's protocol. Briefly, 100.04 ± 0.02 mg of shrimp muscle was mixed with 10 mL of buffer G2 containing 20 μL of 100 mg/mL RNase A (final concentration of 0.2 mg/mL) and 500 μL of 20 mg/mL Proteinase K (final concentration of 1 mg/mL) and incubated at 50 °C for 4 h or until the lysate was clear. Qiagen Genomic-tip 100/G column was equilibrated with five mL of buffer QBT and emptied by gravity flow. The lysate sample was then added into the column and allowed the column to empty by gravity flow. The column was washed twice with 7.5 mL of buffer QC before eluting genomic DNA twice with five mL of 50 °C buffer QF. The eluted genomic DNA was precipitated with four mL of isopropanol, centrifuged at 5,000 g for 15 min at 4 °C. The DNA pellet was washed with five mL of cold 70% ethanol by centrifugation at 5,000 g for 10 min at 4 °C. The DNA pellet was air-dried for 10 min and then resuspended with 100 μL of TE buffer.

E.Z.N.A® Mollusc DNA kit

DNA was extracted following the manufacturer's protocol. Briefly, 50.03 ± 0.03 mg of muscle tissue was mixed with 350 μL of ML buffer. Proteinase K solution (25 μL of 20 mg/mL, final concentration of 1.4 mg/mL) was added to the mixture to digest protein contamination by incubating at 37 °C for 16 h before mixing with one volume of chloroform:isoamyl alcohol (24:1). The reaction was centrifuged at 15,000 g for 10 min at 25 °C. The upper layer of the supernatant was transferred to one volume of BL buffer before 10 μL of 10 mg/mL RNase A were added, incubated at 70 °C for 10 min and cooled-down to room temperature. One volume of 100% ethanol was added and mixed. The HiBind DNA Mini Column was prepared by adding 100 μL of 3M NaOH and centrifuged at 15,000 g for 1 min at 25 °C before 700 μL of sample was transferred to the column. The column was centrifuged and the filtrate was discarded. These steps were repeated until the remaining samples were applied to the column. The column was transferred into a collection tube before adding HBC Buffer (500 μL). The mixture was centrifuged at 13,000 g for 1 min at 25 °C and the filtrate was discarded. The column was washed twice with DNA Wash Buffer (700 μL), followed by centrifugation at 13,000 g for 2 min at 25 °C to dry the membrane. DNA was eluted with the Elution Buffer (50 μL) that had been heated to 70 °C, incubated at room temperature for 5 min and centrifuged at 13,000 g for 2 min at 25 °C.

TIANamp Marine Animals DNA kit

DNA was extracted following the manufacturer's protocol. Briefly, 30.02 ± 0.03 mg of muscle tissue was mixed with 200 μL of Buffer GA by vortexing. RNA was removed by adding 4 μL of 100 mg/mL RNase A (final concentration of 2 mg/mL) at room temperature for 5 min. Proteinase K (20 μL of 20 mg/mL, final concentration of 2 mg/mL) was added and incubated at 56 °C for 3 h. The mixture was added with 200 μL of Buffer GB and incubated at 70 °C for 10 min before 200 μL of 100% ethanol was added. The mixture was transferred into Spin Column CB3 and centrifuged at 13,000 g for 30 s. The column was washed twice with 500 μL of Buffer GD and one time with 600 μL of Buffer PW by centrifugation at 13,000 g for 30 s. The empty column was centrifuged at 13,000 g for 2

min to dry the membrane before DNA was eluted by 50 µL of Buffer TE and centrifuged at 13,000 g for 2 min.

Sbeadex livestock kit

DNA was extracted following the manufacturer's protocol. Briefly, 50.03 ± 0.03 mg of muscle was mixed with 120 µL of PVP buffer, 10 µL of 20 mg/mL protease solution (final concentration of 1.6 mg/mL) and 4.8 µL of debris capture beads and incubated at 55 °C for 2 h. One volume of lysis buffer SB and 5 µL of 20 mg/mL protease solution (final concentration of 0.8 mg/mL) were added to the mixture and incubated at 55 °C for 1 h before centrifugation at 13,000 g at 25 °C for 10 min. The upper aqueous phase was transferred to 180 µL of binding mix and incubated at room temperature for 5 min with constant shaking. The tube was placed in magnetic rack until solution appeared clear, and the supernatant was removed. The beads were washed with 400 µL of wash buffer BN1, wash buffer TN1 and wash buffer TN2, respectively by gently vortexing for 10 min. The tube was placed in a magnetic rack until solution became clear, and the supernatant was removed. The DNA was eluted by 50 µL of Elution buffer AMP and incubated at 55 °C for 10 min. The tube was placed in a magnetic rack until solution appeared clear, and DNA solution was collected to a new tube.

DNA purification

Each genomic DNA sample was cleaned up using AMPure PB bead (Pacific Biosciences, USA) to remove residual genomic DNA isolation reagents prior to the library preparation by adding 0.45 volume of AMPure PB bead to DNA sample. The mixture was incubated by gently vortexing for 10 min, and tubes were placed in magnetic rack until solution appeared clear. The supernatant was removed and washed AMPure PB bead twice with one mL of freshly prepared 70% ethanol. DNA samples were eluted by adding elution buffer and incubated by gently vortexing for 5 min. Tubes were placed in the magnetic rack until solution became clear, and DNA solutions were collected to a new tube.

Assessment of DNA quantity and quality

The extracted DNA by each DNA extraction method was quantified by NanoDrop 8000 spectrophotometer V2.3.2 (Thermo Fisher Science, USA) and Qubit dsDNA BR Assay kit (Invitrogen, USA) using Qubit 2.0 Fluorometer (Invitrogen, USA). The DNA yields of shrimp muscle from different extraction methods were statistically tested using one-way analysis of variance (ANOVA) followed by Duncan's new multiple range test in IBM SPSS statistics 23.0. The DNA quality and integrity were visualized by UV light after electrophoresis in 0.75% of SeaKem® Gold Agarose (Lonza, USA) for pulsed-field gel electrophoresis at 80 volts for 9 h in 0.5x KBB buffer (51 mM Tris, 28 mM TASP, 0.08 mM EDTA, pH 8.7) (Sage Science, USA) containing SYBR Safe DNA gel staining (Invitrogen, USA).

Genomic sequencing analysis

PacBio sequencing was performed at NovogeneAIT (Singapore) following the PacBio's protocol (*Anthony & Kin, 2015*; *Pacific Biosciences of California, 2014*). High-quality HMW

DNA obtained was used to generate a 20-kb SMRTbell library. The SMRTbell library itself was produced by ligating universal hairpin adapters onto double-stranded DNA fragments. The hairpin dimers formed during this process were removed at the end of the protocol using a magnetic bead (AMPure PB) purification step. The final step of the protocol was to remove failed ligation products through the use of exonucleases. After the exonuclease and AMPure PB purification steps, sequencing primer was annealed to the SMRTbell templates following by the binding of polymerase to the annealed templates. According to the effective concentration of library and output requirements, the sample was sequenced on PacBio Sequel platform. The original sequencing reads (called polymerase reads) were processed by the software SMRTlink to filter reads with minimum read quality score of 0.8 and to generate subreads. Each polymerase read was partitioned to form one or more subreads, which contained sequences from a single pass of polymerase on a single strand of an insert within a SMRTbell[TM] template and no adapter sequences. The subreads were used for further analysis.

Data Availability

The genomic data is available in the NCBI under the accession number: PRJNA611030.

RESULTS

DNA quantity

DNA yields were measured by Qubit fluorometer and NanoDrop spectrophotometer (Fig. 1). Among the extraction methods tested, Genomic-tip kit showed the highest genomic DNA yield, while TIANamp and CTAB methods resulted in the lowest genomic DNA yield (Fig. 1).

DNA purity and integrity

Purity of extracted DNA was evaluated based on the ratios between the absorbance values for A260/280 and A260/230 using a NanoDrop spectrophotometer (Table 1). Recommended purity of DNA for a long-read sequencing platform is 1.8–2.0 for A260/280 ratio and ∼2.0 for A260/230 ratio (Hassan et al., 2015; Lucena-Aguilar et al., 2016; Usman et al., 2014). Mollusc and Genomic-tip kits were the only two methods that yielded DNA with the purity that passed the criteria for tissue sample. On the other hand, the CTAB protocol and TIANamp kit did not yield DNA samples that passed the criteria in any of the samples tested. DNA extracted using the Sbeadex livestock kit passed the criteria in A260/280, but not A260/A230 ratio (Table 1).

Integrity of genomic DNA samples was evaluated using pulsed-field gel electrophoresis of the 100 ng of DNA samples obtained using five DNA extraction methods after the AMPure PB bead purification step (Fig. S1). However, DNA extracted by Mollusc kit (Fig. S1C) and Sbeadex livestock kit (Fig. S1E) could not be clearly visualized; therefore, 200 ng of DNA samples from these two methods were used to improve visualization allowing better comparison of the DNA integrity in Fig. 2. DNA extracted from muscle tissue by CTAB method and Genomic-tip kit appeared to have the highest molecular weight (Figs. 2A and 2B), while the Mollusc DNA Kit, TIANamp Marine Animals DNA kit and Sbeadex
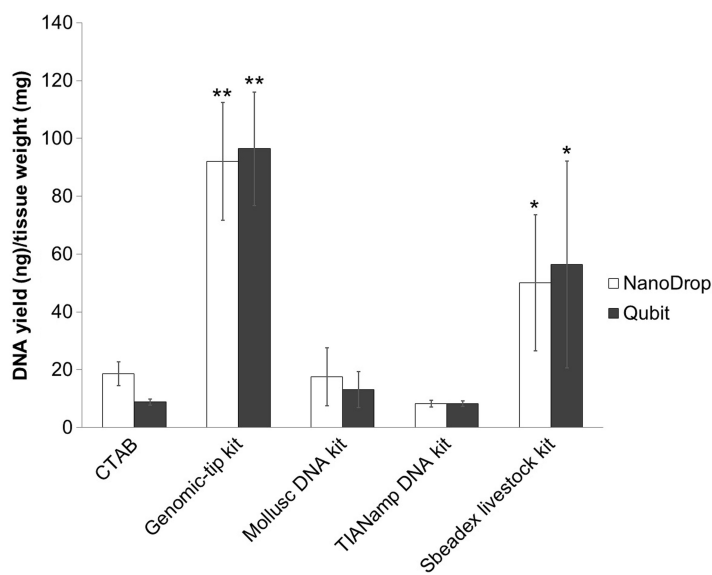
Figure 1 **Average DNA yield measured with NanoDrop spectrophotometer and Qubit Fluorometer of shrimp muscle using different DNA extraction methods after purified by AMPure PB bead.** Error bars indicate standard deviation of the mean from four replicates. The asterisk indicates significantly different level among the extraction methods: * $P < 0.05$ and ** $P < 0.01$.

Table 1 **Quality of gDNA extracted from shrimp muscle using different DNA extraction methods after purified with AMPure PB bead represented by an average ± a standard deviation value (SD).**

NanoDrop spectrophotometer	Extraction method				
	CTAB	Genomic-tip 100/G kit	Mollusc DNA kit	TIANamp Marine Animals DNA kit	Sbeadex livestock kit
A260/A280	1.40 ± 0.09	1.84 ± 0.01	1.98 ± 0.04	1.63 ± 0.16	1.81 ± 0.03
A260/A230	0.40 ± 0.05	2.45 ± 0.03	2.01 ± 0.16	1.74 ± 0.22	1.85 ± 0.36

livestock kit yielded DNA with smeary bands, indicating shearing of DNA (Figs. 2C, 2D and 2E).

Sequencing quality

To evaluate whether the quality of the obtained genomic DNA was sufficiently high for the long-read sequencing technology, the genomic DNA extracted by the top three extraction methods (Genomic-tip kit, CTAB method and Mollusc DNA kit) were used to prepare 20Kb libraries and subsequently sequenced using the PacBio (Fig. 3). While the genomic DNA from Genomic-tip and Mollusc DNA kits allowed successful library construction, the one from CTAB method did not. The sequencing of DNA obtained from Genomic-tip kit yielded 5.36 Gb/SMRT cell with the longest length of 272.89 Kb, an average length of 10.28 Kb and an N50 length of 14.57 Kb. The number of reads longer than 20 Kb represented 11.44% of the total reads. The sequencing of DNA obtained from Mollusc DNA kit gave 3.07 Gb/ SMRT cell with the maximum read length of 112.41 Kb, an average length of
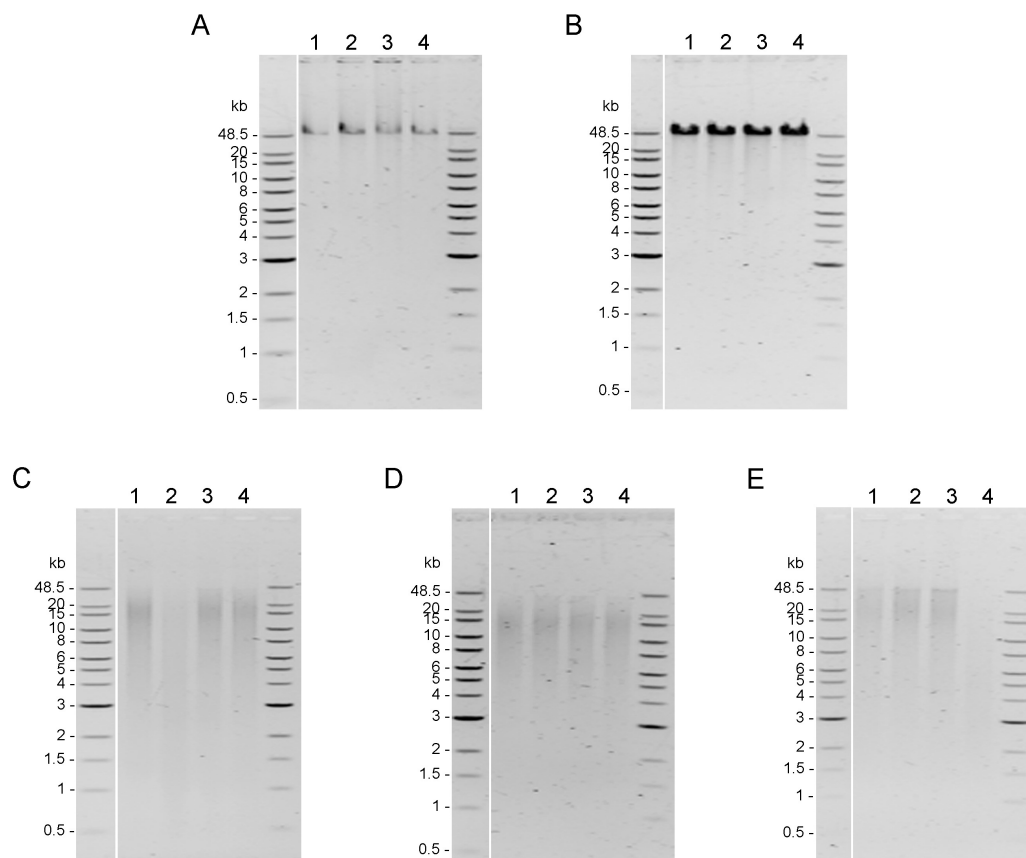
Figure 2 Quality assessment of genomic DNA of shrimp muscle extracted with different DNA extraction methods. After purified by AMPure PB bead, gDNA (n=4) extracted from (A) CTAB method (100 ng of gDNA), (B) QIAGEN Genomic-tip 100/G kit (100 ng of gDNA), (C) E.Z.N.A.®Mollusc DNA Kit (200 ng of gDNA), (D) TIANamp Marine Animals DNA kit (100 ng of gDNA) and (E) Sbeadex livestock kit (200 ng of gDNA) were loaded on 0.75% pulsed-field gel electrophoresis and run at 80 Volts for 9 h. The DNA size marker is Quick-Load 1 kb Extend DNA Ladder (New England BioLabs).

4.40 Kb and an N50 length of 9.74 Kb. The reads longer than 20 Kb covered 2.42% of total reads.

DISCUSSION

Availability of genome sequences for various organisms has revolutionized biology. Recent advances in the long-read sequencing technology have promised even faster and better quality of genome sequences. With such power of the recent sequencing technology, the Earth BioGenome Project has announced to sequence, catalog and characterize the genomes of Earth's eukaryotic biodiversity over a period of ten years (*Lewin et al., 2018*). While the capacity of the sequencing technology theoretically can make such a grand challenge possible, numerous organisms are intrinsically difficult to be sequenced.

Marine invertebrates are deemed to be ones of the most challenging species to be sequenced. This is possibly due to their DNA contents containing high polysaccharide
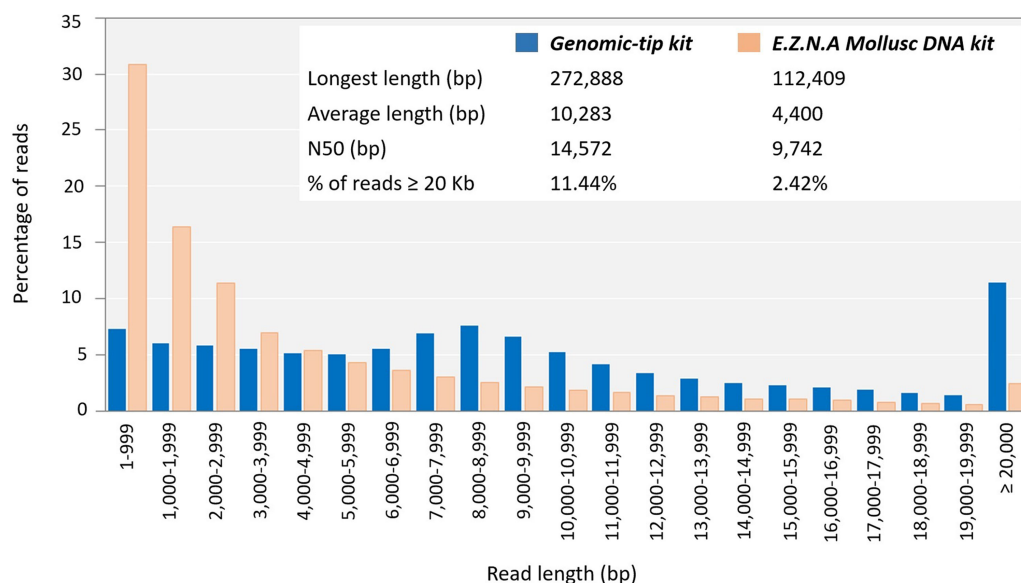
Figure 3 **Read length distribution using DNA extracted by Mollusc DNA kit and Genomic-tip 100/G kit.** Mollusc DNA kit was sequenced for two SMRT cells and Genomic-tip 100/G kit was sequenced for 37 SMRT cells.

content, polyphenols and other secondary metabolites which can interfere downstream DNA library preparations (*Panova et al., 2016*). Some of the genomes of marine organisms such as water flea (*Daphnia pulex*) (*Ye et al., 2017*), Atlantic cod (*Gadus morhua*) (*Star et al., 2016*), black tiger shrimp (*Penaeus monodon*) (*Uengwetwanit et al., 2020*; *Yuan et al., 2018*) and Pacific white shrimp (*Litopenaeus vannamei*) (*Zhang et al., 2019*) reportedly contain high density of repetitive sequences, which hinder high quality genome sequence assembly. To overcome this challenge, high quality and integrity of the starting DNA must be used to fully exploit the capacity of the long-read sequencing platform. Thus, this study embarks on comparing the efficacy of five already established protocols for genomic DNA extraction of a marine organism using the black tiger shrimp as a model.

Among the five protocols examined, CTAB method is the only one where solutions/buffers required can be prepared in-house, hence, has the advantage of being an inexpensive and simple approach for DNA extraction (Table 2). It has been successfully used to extract DNA from organisms containing a high level of polysaccharide such as plant samples (*Abdel-Latif & Osman, 2017*; *Doyle & Doyle, 1987*; *Michiels et al., 2003*). CTAB method also utilizes a reducing agent β-mercaptoethanol in lysate solution, which prevents oxidative damage of nucleic acid, resulting in HMW DNA in other organisms (*Bitencourt et al., 2007*; *Herzer, 2001*). In this study, although the DNA obtained from CTAB method remained at high molecular weight as expected, it yielded the lowest purity of DNA sample based on A260/A280 and A260/A230 ratios (Table 2). The low ratios obtained suggest ineffective removal of proteins and organic compounds. This is not surprising because the standard CTAB method does not require proteinase K treatment but uses phenol/chloroform extraction steps, which introduces organic contaminants and

Table 2 Performance comparison of DNA extraction methods.

	CTAB	Genomic-tip	Mollusc DNA	TIANamp Marine Animals DNA	Sbeadex livestock
Time					
-Total assay time	2 h 30 min	7-8 h	17 h (overnight lysis step)	3 h	4 h
Hands-on time	1 h 30 min	1 h 30 min	1 h	45 min	2 h
Cost[a]	Inexpensive (2 US-D/rnx/100 mg tissue)	Expensive (27.4 US-D/rnx/100 mg tissue)	Expensive (8.2 US-D/rnx/50 mg tissue)	Inexpensive (2.5 US-D/rnx/30 mg tissue)	Expensive (7.2 US-D/rnx/50 mg tissue)
Yield	Low yield (8.81 ng DNA/mg tissue)	Highest yield (96.38 ng DNA/mg tissue)	Low yield (13.15 ng DNA/mg tissue)	Lowest yield (8.25 ng DNA/mg tissue)	Moderate yield (56.42 ng DNA/mg tissue)
Quality					
— Integrity	HMW observed in >48.5 kb	HMW observed in >48.5 kb	Smearing of DNA observed in <20 kb	Smearing of DNA observed in <20 kb	HMW observed in 48.8 kb and smearing of DNA observed in <48.5 kb
— Purity	Poor	Good	Good	Poor	Poor

Notes.

[a]Cost indicated here was converted from the local currency to USD. It may vary in other countries depending on costs from local distributors.

causes yield loss. To further improve this method for better quality of DNA, a treatment of proteinase K or other purification steps should be considered.

The other four extraction methods are commercial kits. Genomic-tip kit employs gravity-flow, anion-exchange tips to minimize DNA shearing from centrifugation, vortexing, and pipetting steps. This gravity-flow based method gave the highest yield when compared with other methods (Table 2). This method simultaneously lyses and enzymatically digests proteins in the sample, allowing immediate denaturation of proteins such as nucleases, histones, and DNA-binding proteins (*Qiagen, 2015*). These proteins have been reported to affect the quality of DNA (*Chacon-Cortes & Griffiths, 2014*; *Tan & Yiap, 2009*). Shrimp muscle consists of numerous fiber proteins, majorly composed of myofibrillar and sarcoplasmic proteins (*Medler & Mykles, 2015*). Thus, digesting these fibrous proteins in shrimp muscle using proteinase K from the beginning may have resulted in the highest DNA yield among the extraction methods. Other fibrous samples have witnessed the same high yield benefit from performing proteinase K digestion at the beginning of the protocol as well (*Vaidya et al., 2018*). In addition to the high yield, the highest integrity of the DNA obtained from this kit was also evidenced by the longest read length from the sequencing result.

Mollusc kit utilizes cationic detergent, combined with polysaccharide property of CTAB and the selective DNA binding matrix. This DNA extraction kit was used for muscle of marine metazoans to remove polysaccharide contamination (*Lobo et al., 2013*). Using this kit with shrimp tissue gave high purity of DNA; however, the pulsed-field gel electrophoresis suggested that the DNA was sheared. The sequencing result also confirmed that the DNA from the kit was not intact as the average read length was much shorter than those from Genomic-tip kit. This kit requires the most number of centrifugation steps (7 times) when compared with other methods (CTAB method, 6 times; Genomic-tip kit, 2 times;

TIANamp kit, 6 times and Sbeadex livestock kit, 1 time), which might cause the shearing of DNA (*McDonald, 2002*; *Reigstad, Bartossek & Schleper, 2011*; *Sage Science, 2020*).

For TIANamp kit, the mechanism for DNA extraction is very similar to that of Mollusc kit without CTAB and chloroform extraction step. The DNA obtained from this kit was one of the lowest quality and purity among the examined commercial kits. This result was indeed unexpected as the kit was successfully employed to extract genomic DNA of the Pacific white shrimp (*L. vannamei)* enabling a complete draft of genome sequence using a PacBio sequencing technology (*Zhang et al., 2019*). Given no details of the protocol from the paper, it might be possible that some modifications of the DNA extraction protocol were made to obtain high quality and intact DNA sufficient to give successful sequencing results.

Unlike the other commercial kits examined in this study, Sbeadex livestock kit utilizes paramagnetic microparticles whose surface binds DNA through an anion-exchange mechanism. The DNA obtained from this bead-based kit was of high purity in most samples but the integrity was compromised as observed in the pulsed-field gel electrophoresis. Due to the absence of RNaseA treatment step in this protocol, the DNA integrity might be jeopardized. Indeed, RNaseA treatment was reportedly necessary for isolation of high quality genomic DNA (*Healey et al., 2014*; *Lienhard & Schäffer, 2019*; *Rana et al., 2019*).

It is important to note that this study employed the established protocol for each method without further modifications. Indeed, each method can be further modified to give better extraction efficacy, resulting in better quality and yield. For example, reducing incubation time of proteinase K, which can cause DNA degradation (*Athanasio et al., 2016*), might result in better integrity of DNA. Optimizing animal tissues for each method can also improve digestion of the tissue to release more DNA yield (*Jasrotia & Langer, 2019*). The inclusion of RNaseA treatment can alleviate the issue of DNA degradation (*Healey et al., 2014*).

To further demonstrate the suitability for long-read sequencing of the genomic DNA obtained in this study, the DNA extracted by the top three methods (Genomic-tip kit, CTAB method, and Mollusc DNA kit) were subsequently sequenced by the PacBio technology. Genomic-tip and Mollusc DNA kits were only two methods that allowed successful library construction. This indicates that both high-molecular weight and purity of genomic DNA were equally important to successful long-read sequencing. Purity is essential to generate sequencing library, whereas the high-molecular weight of genomic DNA is critical for construction of large insert library. Despite highly intact DNA, CTAB method is not efficient at removing contaminants. Therefore, the protocol might need to be optimized for particular organism to compensate purity and quality (*Arseneau, Steeves & Laflamme, 2017*; *Chakraborty, Saha & Ananthram, 2020*).

As a matter of fact, the Mollusc kit employed in this study was a CTAB-based method that has been optimized for mollusk. However, when used with shrimp sample, it only showed improvement of purity but not integrity. Degradation of DNA might be due to long exposure of proteinase K (*Athanasio et al., 2016*). Mollusc kit used overnight proteinase K and RNaseA digestion, whereas CTAB method required a short incubation with RNaseA (<30 min) and Genomic-tip kit used 4 h incubation of proteinase K and RNaseA. Long

incubation periods for proteinase K digestion in Mollusc kits resulted in fragmented DNA as it could be observed that 2.42% of reads were remained in read length $\geq$ 20 Kb in comparison to 11.44% of read length $\geq$ 20 Kb remaining in Genomic-tip kit. Beside longer sequencing reads (14.57 Kb of N50), Genomic-tip kit generated higher sequencing yield (5.36 Gb/ SMRT cell). Genomic-tip kit purifies DNA based on gravity-flow and anion-exchange tip, which are less prone to shearing DNA compared to vortexing, pipetting and long exposure of enzyme digestion. As demonstrated in the present study, Genomic-tip kit is a suitable method for long-read sequencing. Although Genomic-tip kit is relatively time-consuming and expensive, the high-quality and yield of sequencing reads outweigh these disadvantages (Table 2). However, for laboratories with limited resources and access, further modifications of less expensive methods such as CTAB exaction protocol might be considered.

CONCLUSIONS

The quality of starting DNA for a long-read sequencing platform such as PacBio is a key determinant of the successful result. In this study, the efficiency of extracting high-quality HMW DNA was evaluated for five established DNA extraction methods without any modifications. It is noteworthy that further modifications of the established protocols might yield even better DNA quality and quantity. The gravity-flow based method was found to yield the highest quality genomic DNA and hence suitable for downstream whole genome sequencing application using the long-read sequencing platform. This protocol could be further used for extracting high quality genomic DNA for long-read sequencing of other marine organisms, allowing full exploitation of the sequencing platform.

ACKNOWLEDGEMENTS

We are grateful to Prof. Morakot Tanticharoen (NSTDA), Dr. Kanyawim Kirtikara (BIOTEC), and Dr. Wonnop Visessanguan (BIOTEC) for their mentorship on the shrimp research programs. We are thankful to Prof. Frédéric Tangy (Institut Pasteur) and Dr. Rungnapa Leelatanawit (BIOTEC) for technical advice.

Funding

This project was funded by National Center for Genetic Engineering and Biotechnology (BIOTEC Platform No. P1651718), National Science and Technology Development Agency (TDG Food and Feed program No. P1950419) and supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 734486 (SAFE-Aqua). National Institute of General Medical Sciences of the National Institutes of Health (P20GM125503) awarded to Dr. Intawat Nookaew. There was no additional external funding received for this study. Visiting Professor Program 2019 was awarded to Dr. Intawat Nookaew by the National Science and Technology

Development Agency, Thailand. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

National Center for Genetic Engineering and Biotechnology (Platform): P1651718.

National Science and Technology Development Agency (TDG Food and Feed program): P1950419.

The European Union's Horizon 2020 research and innovation programme under the Marie Skᵃodowska-Curie (SAFE-Aqua): 734486.

National Science and Technology Development Agency (Visiting Professor Program 2019). National Institute of General Medical Sciences of the National Institutes of Health: P20GM125503.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Pacharaporn Angthong conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Tanaporn Uengwetwanit performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Wirulda Pootakham, Intawat Nookaew and Nitsara Karoonuthaisiri conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Kanchana Sittikankaew, Duangjai Sangsrakru and Thippawan Yoocha performed the experiments, prepared figures and/or tables, and approved the final draft.
- Chutima Sonthirod performed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Thidathip Wongsurawat and Piroon Jenjaroenpun analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Wanilada Rungrassamee conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The data is available at NCBI: PRJNA611030 and at the Black Tiger Shrimp (*Penaeus monodon*) Genome Project (raw genomic PacBio reads which used DNA obtained from Genomic-tip kit and Mollusc DNA kit).

Link "Black Tiger Shrimp (*Penaeus monodon*) Genome Project" to http://www.biotec.or.th/pmonodon/index.php/downloads.

PeerJ —————————————————————————————

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.10340#supplemental-information.

REFERENCES

Abdel-Latif A, Osman G. 2017. Comparison of three genomic DNA extraction methods to obtain high DNA quality from maize. *Plant Methods* **13**:1 DOI 10.1186/s13007-016-0152-4.

Amarasinghe SL, Su S, Dong X, Zappia L, Ritchie ME, Gouil Q. 2020. Opportunities and challenges in long-read sequencing data analysis. *Genome Biology* **21**:30 DOI 10.1186/s13059-020-1935-5.

Anderson JL, Asche F, Garlock T, Chu J. 2017. Aquaculture: its role in the future of food. In: *World agricultural resources and food security. Volume 17.* Bingley: Emerald Publishing Limited, 159–173.

Anthony R, Kin FA. 2015. PacBio sequencing and its applications, genomics proteomics. *Bioinformatics* **13**:278–289 DOI 10.1016/j.gpb.2015.08.002.

Arseneau JR, Steeves R, Laflamme M. 2017. Modified low-salt CTAB extraction of high-quality DNA from contaminant-rich tissues. *Molecular Ecology Resources* **17**:686–693 DOI 10.1111/1755-0998.12616.

Athanasio CG, Chipman JK, Viant MR, Mirbahai L. 2016. Optimisation of DNA extraction from the crustacean *Daphnia. PeerJ* **4**:e2004–e2004 DOI 10.7717/peerj.2004.

Berthelot C, Brunet F, Chalopin D, Juanchich A, Bernard M, Noël B, Bento P, Da Silva C, Labadie K, Alberti A, Aury J-M, Louis A, Dehais P, Bardou P, Montfort J, Klopp C, Cabau C, Gaspin C, Thorgaard GH, Boussaha M, Quillet E, Guyomard R, Galiana D, Bobe J, Volff J-N, Genêt C, Wincker P, Jaillon O, Crollius HR, Guiguen Y. 2014. The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nature Communications* **5**:3657 DOI 10.1038/ncomms4657.

Bitencourt JVT, Roratto PA, Bartholomei-Santos ML, Santos S. 2007. Comparison of different methodologies for DNA extraction from *Aegla longirostri. Brazilian Archives of Biology and Technology* **50**:989–994 DOI 10.1590/S1516-89132007000700010.

Chacon-Cortes D, Griffiths L. 2014. Methods for extracting genomic DNA from whole blood samples: current perspectives. *Journal of Biorepository Science for Applied Medicine* **2**:1–9 DOI 10.2147/BSAM.S46573.

Chakraborty S, Saha A, Neelavar Ananthram A. 2020. Comparison of DNA extraction methods for non-marine molluscs: is modified CTAB DNA extraction method more efficient than DNA extraction kits? *3 Biotech* **10**:69 DOI 10.1007/s13205-020-2051-7.

Collins FS, Green ED, Guttmacher AE, Guyer MS. 2003. A vision for the future of genomics research. *Nature* **422**:835–847 DOI 10.1038/nature01626.

Davidson WS, Koop BF, Jones SJM, Iturra P, Vidal R, Maass A, Jonassen I, Lien S, Omholt SW. 2010. Sequencing the genome of the Atlantic salmon (*Salmo salar*). *Genome Biology* **11**:403 DOI 10.1186/gb-2010-11-9-403.

Doyle JJ, Doyle JJ. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**:11–15.

FAO. 2016. *Food and Agriculture Organization of the United Nations Statistics. Fisheries and aquaculture software. FishStat Plus - Universal software for fishery statistical time series.* Rome: Fisheries and Aquaculture Department.

Geraci C, Varzandi AR, Schiavo G, Bovo S, Ribani A, Utzeri VJ, Galimberti G, Buttaz-zoni L, Ovilo C, Gallo M, Dall'Olio S, Fontanesi L. 2019. Genetic markers associated with resistance to infectious diseases have no effects on production traits and haematological parameters in Italian Large White pigs. *Livestock Science* **223**:32–38 DOI 10.1016/j.livsci.2019.03.003.

Guppy JL, Jones DB, Jerry DR, Wade NM, Raadsma HW, Huerlimann R, Zenger KR. 2018. The state of ''omics'' research for farmed penaeids: advances in research and impediments to nndustry utilization. *Frontiers in Genetics* **9**:282 DOI 10.3389/fgene.2018.00282.

Hassan R, Husin A, Sulong S, Yusoff S, Johan MF, Yahaya BH, Ang CY, Ghazali S, Cheong SK. 2015. Guidelines for nucleic acid detection and analysis in hematological disorders. *Malaysian Journal of Pathology* **37**:165–173.

Healey A, Furtado A, Cooper T, Henry RJ. 2014. Protocol: a simple method for extracting next-generation sequencing quality genomic DNA from recalcitrant plant species. *Plant Methods* **10**:21 DOI 10.1186/1746-4811-10-21.

Herzer S. 2001. DNA purification. In: *Molecular biology problem solver: a laboratory guide.* New York: Wiley-Liss, 167–196.

Jasrotia R, Langer S. 2019. Comparative account of DNA extraction protocols in some freshwater prawns of Genus *Macrobrachium* (Bate, 1868) (Family Palaemonidae) from Jammu waters for PCR based applications. *Biomedical and Pharmacology Journal* **12**:1201–1206 DOI 10.13005/bpj/1749.

Lewin HA, Robinson GE, Kress WJ, Baker WJ, Coddington J, Crandall KA, Durbin R, Edwards SV, Forest F, Gilbert MTP, Goldstein MM, Grigoriev IV, Hackett KJ, Haussler D, Jarvis ED, Johnson WE, Patrinos A, Richards S, Castilla-Rubio JC, Van Sluys M-A, Soltis PS, Xu X, Yang H, Zhang G. 2018. Earth BioGenome Project: Sequencing life for the future of life. *Proceedings of the National Academy of Sciences of the United States of America* **115**:4325–4333 DOI 10.1073/pnas.1720115115.

Li DF, Lian L, Qu LJ, Chen YM, Liu WB, Chen SR, Zheng JX, Xu GY, Yang N. 2013. A genome-wide SNP scan reveals two loci associated with the chicken resistance to Marek's disease. *Animal Genetics* **44**:217–222 DOI 10.1111/j.1365-2052.2012.02395.x.

Lienhard A, Schäffer S. 2019. Extracting the invisible: obtaining high quality DNA is a challenging task in small arthropods. *PeerJ* **7**:e6753 DOI 10.7717/peerj.6753.

Lobo J, Costa PM, Teixeira MAL, Ferreira MSG, Costa MH, Costa FO. 2013. Enhanced primers for amplification of DNA barcodes from a broad range of marine metazoans. *BMC Ecology* **13**:34 DOI 10.1186/1472-6785-13-34.
Lucena-Aguilar G, Sánchez-López AM, Barberán-Aceituno C, Carrillo-Ávila JA, López-Guerrero JA, Aguilar-Quesada R. 2016. DNA source selection for downstream applications based on DNA quality indicators analysis. *Biopreservation and Biobanking* **14**:264–270 DOI 10.1089/bio.2015.0064.

Martínez R, Dunner S, Toro R, Tobón J, Gallego J, Cañón J. 2010. Effect of polymorphisms in the Slc11a1 coding region on resistance to brucellosis by macrophages in vitro and after challenge in two Bos breeds (Blanco Orejinegro and Zebu). *Genetics and Molecular Biology* **33**:463–470 DOI 10.1590/S1415-47572010000300014.

Martínez R, Rocha J, Bejarano D, Gomez Y, Abuabara Y, Gallego J. 2016. Identification of SNPs in growth-related genes in Colombian creole cattle. *Genetics and Molecular Research* **15**:15038762 DOI 10.4238/gmr.15038762.

McDonald JH. 2002. CTAB protocol for mollusc DNA. *Available at http:// udel.edu/ ~mcdonald/ ctab.html*.

Medler S, Mykles D. 2015. Natural history of crustacea. In: Chang ES, Thiel M, eds. *Chapter: Muscle structure, fiber types, and physiology.* Oxford: Oxford University Press, 103–133.

Metzker ML. 2010. Sequencing technologies—the next generation. *Nature Reviews Genetics* **11**:31–46 DOI 10.1038/nrg2626.

Michiels A, Dew Van, Tucker M, Van RL, Van LA. 2003. Extraction of high-quality genomic DNA from latex-containing plants. *Analytical Biochemistry* **315**:85–89 DOI 10.1016/S0003-2697(02)00665-6.

Nadia B, Sadok S. 2017. Comparison of DNA-extraction methods suitable for PCR-based applications to identify shrimp species in commercial products. *Journal of FisheriesSciences com* **11**:25–31 DOI 10.21767/1307-234X.1000138.

Pacific Biosciences of California. 2014. *Template preparation and sequencing guide.* Menlo Park: Pacific Biosciences of California, Inc.

Panova M, Aronsson H, Cameron R, Dahl P, Godhe A, Lind U, Ortega-Martinez O, Pereyra R, Tesson S, Wrange A-L, Blomberg A, Johannesson K. 2016. DNA extraction protocols for whole-genome sequencing in marine organisms. *Methods in Molecular Biology* **1452**:13–44.

Pereira JC, Chaves R, Bastos E, Leitão A, Guedes-Pinto H. 2011. An efficient method for genomic DNA extraction from different molluscs species. *International Journal of Molecular Sciences* **12**:8086–8095 DOI 10.3390/ijms12118086.

Pollard MO, Gurdasani D, Mentzer AJ, Porter T, Sandhu MS. 2018. Long reads: their purpose and place. *Human Molecular Genetics* **27**:234–241 DOI 10.1093/hmg/ddy177.

Psifidi A, Dovas CI, Bramis G, Lazou T, Russel CL, Arsenos G, Banos G. 2015. Comparison of eleven methods for genomic DNA extraction suitable for large-scale whole-genome genotyping and long-term DNA banking using blood samples. *PLOS ONE* **10**:e0115960 DOI 10.1371/journal.pone.0115960.

Qiagen. 2015. Protocol: preparation of tissue samples. In: *Qiagen genomic DNA handbook 06/2015.* Hilden: Qiagen, 31–36.

Rahman S, Schmidt D, Hughes J. 2017. Development of an efficient method for producing high quality genomic DNA in crustaceans. *Asian Journal of Animal Sciences* **11**:214–220 DOI 10.3923/ajas.2017.214.220.

Ramanan VK, Shen L, Moore JH, Saykin AJ. 2012. Pathway analysis of genomic data: concepts, methods, and prospects for future development. *Trends in Genetics* **28**:323–332 DOI 10.1016/j.tig.2012.03.004.

Rana MM, Aycan M, Takamatsu T, Kaneko K, Mitsui T, Itoh K. 2019. Optimized nuclear pellet method for extracting next-generation sequencing quality genomic DNA from fresh leaf tissue. *Methods and Protocols* **2**:54 DOI 10.3390/mps2020054.

Reigstad LJ, Bartossek R, Schleper C. 2011. Chapter thirteen - preparation of high-molecular weight DNA and metagenomic libraries from soils and hot springs. In: Klotz MG, Stein LY, eds. *Methods in enzymology.* New York: Academic Press, 319–344.

Sage Science. 2020. *How to isolate DNA from mitochondria: an energetic guide.* Available at https://bitesizebio.com/47040/isolate-dna-mitochondria/ .

Schadt EE, Turner S, Kasarskis A. 2010. A window into third-generation sequencing. *Human Molecular Genetics* **19**:R227–R240 DOI 10.1093/hmg/ddq416.

Star B, Hansen MHS, Skage M, Bradbury IR, Godiksen JA, Kjesbu OS, Jentoft S. 2016. Preferential amplification of repetitive DNA during whole genome sequencing library creation from historic samples. *Science and Technology of Archaeological Research* **2**:36–45 DOI 10.1080/20548923.2016.1160594.

Star B, Nederbragt AJ, Jentoft S, Grimholt U, Malmstrøm M, Gregers TF, Rounge TB, Paulsen J, Solbakken MH, Sharma A, Wetten OF, Lanzén A, Winer R, Knight J, Vogel J-H, Aken B, Andersen Ø, Lagesen K, Tooming-Klunderud A, Edvardsen RB, Tina KG, Espelund M, Nepal C, Previti C, Karlsen BO, Moum T, Skage M, Berg PR, Gjøen T, Kuhl H, Thorsen J, Malde K, Reinhardt R, Du L, Johansen SD, Searle S, Lien S, Nilsen F, Jonassen I, Omholt SW, Stenseth NC, Jakobsen KS. 2011. The genome sequence of Atlantic cod reveals a unique immune system. *Nature* **477**:207–210 DOI 10.1038/nature10342.

Tan SC, Yiap BC. 2009. DNA, RNA, and protein extraction: the past and the present. *Journal of Biomedicine and Biotechnology* **2009**:574398 DOI 10.1155/2009/574398.

Uengwetwanit T, Pootakham W, Nookaew I, Sonthirod C, Angthong P, Sittikankaew K, Rungrassamee W, Arayamethakorn S, Wongsurawat T, Jenjaroenpun P, Sangsrakru D, Leelatanawit R, Khudet J, Koehorst JJ, Schaap PJ, Dos Santos VM, Tangy F, Karoonuthaisiri N. 2020. A chromosome-level assembly of the black tiger shrimp (*Penaeus monodon*) genome facilitates the identification of novel growth-associated genes. *bioRxiv* DOI 10.1101/2020.05.14.096073.

Usman T, Yu Y, Liu C, Fan Z, Wang Y. 2014. Comparison of methods for high quantity and quality genomic DNA extraction from raw cow milk. *Genetics and Molecular Research* **13**:3319–3328 DOI 10.4238/2014.April.29.10.

Vaidya JD, Van den Bogert B, Edwards JE, Boekhorst J, Van Gastelen S, Saccenti E, Plugge CM, Smidt H. 2018. The effect of DNA extraction methods on observed

microbial communities from fibrous and liquid rumen fractions of dairy cows. *Frontiers in Microbiology* **9**:92 DOI 10.3389/fmicb.2018.00092.

Winnepenninckx B, Backeljau T, De Wachte R. 1993. Extraction of high molecular weight DNA from molluscs. *Trends in Genetics* **9**:407 DOI 10.1016/0168-9525(93)90102-n.

Xie L, Luo C, Zhang C, Zhang R, Tang J, Nie Q, Ma L, Hu X, Li N, Da Y, Zhang X. 2012. Genome-wide association study identified a narrow chromosome 1 region associated with chicken growth traits. *PLOS ONE* **7**:e30910 DOI 10.1371/journal.pone.0030910.

Yang Q, Wu P, Wang K, Chen D, Zhou J, Ma J, Li M, Xiao W, Jiang A, Jiang Y, Bai L, Zhu L, Li X, Tang G. 2019. SNPs associated with body weight and backfat thickness in two pig breeds identified by a genome-wide association study. *Genomics* **111**:1583–1589 DOI 10.1016/j.ygeno.2018.11.002.

Ye Z, Xu S, Spitze K, Asselman J, Jiang X, Ackerman MS, Lopez J, Harker B, Raborn RT, Thomas WK, Ramsdell J, Pfrender ME, Lynch M. 2017. A new reference genome assembly for the microcrustacean *Daphnia pulex*. *G3* **7**:1405–1416 DOI 10.1534/g3.116.038638.

Yuan J, Zhang X, Liu C, Yu Y, Wei J, Li F, Xiang J. 2018. Genomic resources and comparative analyses of two economical penaeid shrimp species, *Marsupenaeus japonicus* and *Penaeus monodon*. *Marine Genomics* **39**:22–25 DOI 10.1016/j.margen.2017.12.006.

Zhang G, Fang X, Guo X, Li L, Luo R, Xu F, Yang P, Zhang L, Wang X, Qi H, Xiong Z, Que H, Xie Y, Holland PWH, Paps J, Zhu Y, Wu F, Chen Y, Wang J, Peng C, Meng J, Yang L, Liu J, Wen B, Zhang N, Huang Z, Zhu Q, Feng Y, Mount A, Hedgecock D, Xu Z, Liu Y, Domazet-Lošo T, Du Y, Sun X, Zhang S, Liu B, Cheng P, Jiang X, Li J, Fan D, Wang W, Fu W, Wang T, Wang B, Zhang J, Peng Z, Li Y, Li N, Wang J, Chen M, He Y, Tan F, Song X, Zheng Q, Huang R, Yang H, Du X, Chen L, Yang M, Gaffney PM, Wang S, Luo L, She Z, Ming Y, Huang W, Zhang S, Huang B, Zhang Y, Qu T, Ni P, Miao G, Wang J, Wang Q, Steinberg CEW, Wang H, Li N, Qian L, Zhang G, Li Y, Yang H, Liu X, Wang J, Yin Y, Wang J. 2012. The oyster genome reveals stress adaptation and complexity of shell formation. *Nature* **490**:49–54 DOI 10.1038/nature11413.

Zhang X, Yuan J, Sun Y, Li S, Gao Y, Yu Y, Liu C, Wang Q, Lv X, Zhang X, Ma KY, Wang X, Lin W, Wang L, Zhu X, Zhang C, Zhang J, Jin S, Yu K, Kong J, Xu P, Chen J, Zhang H, Sorgeloos P, Sagi A, Alcivar-Warren A, Liu Z, Wang L, Ruan J, Chu KH, Liu B, Li F, Xiang J. 2019. Penaeid shrimp genome provides insights into benthic adaptation and frequent molting. *Nature Communications* **10**:356 DOI 10.1038/s41467-018-08197-4.

Zwane AA, Schnabel RD, Hoff J, Choudhury A, Makgahlela ML, Maiwashe A, Van Marle-Koster E, Taylor JF. 2019. Genome-wide SNP discovery in indigenous cattle breeds of South Africa. *Frontiers in Genetics* **10**:273 DOI 10.3389/fgene.2019.00273.