A peer-reviewed version of this preprint was published in PeerJ on 11 October 2016.

Mulcahy DG, Macdonald KS III, Brady SG, Meyer C, Barker KB, Coddington J. 2016. Greater than *X* kb: a quantitative assessment of preservation conditions on genomic DNA quality, and a proposed standard for genome-quality DNA. PeerJ 4:e2528
https://doi.org/10.7717/peerj.2528

Greater than *X* kb: A quantitative assessment of preservation conditions on genomic DNA quality, and a proposed standard for genome-quality DNA

Daniel G. Mulcahy [Corresp., 1], **Kenneth S. Macdonald III** [2], **Seán G. Brady** [1], **Christopher Meyer** [3], **Katharine B. Barker** [1], **Jonathan Coddington** [1, 4]

[1] Global Genome Initiative, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA
[2] Laboratories of Analytical Biology, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA
[3] Department of Invertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA
[4] Department of Entomology, National Museum of Natural History, Smithsonian Institution, Washington, DC, United States

Corresponding Author: Daniel G. Mulcahy
Email address: MulcahyD@si.edu

Advances in biodiversity genomic sequencing will increasingly depend on the availability of DNA samples—and their quantifiable metadata—preserved in large institutional biorepositories that are discoverable to the scientific community. Improvements in sequencing technology constantly provide longer reads, such that longer fragment length, higher molecular weight, and overall “genome-quality” DNA (gDNA) will be desirable. Ideally, biorepositories should publish numerical scale measurements of DNA quality useful to the user community. However, the most widely used technique to evaluate DNA quality, the classic agarose gel, has yet to be quantified. Here we propose a simple and economical method using open source image analysis software to make gDNA gel images quantifiable, and propose percentage of gDNA “greater than *X* kb” as a standard of comparison, where *X* is a band from any widely used DNA ladder with desirably large band sizes. We employ two metadata standards (“DNA Threshold” and “Percent above Threshold”) introduced as part of the Global Genome Biodiversity Network (GGBN) Darwin Core extension. We illustrate the method using the traditionally used *Hin*dIII ladder and the 9,416 base-pair (bp) band as a standard. We also present data, for two taxa, a vertebrate (fish) and an invertebrate (crab), on how gDNA quality varies with seven tissue preservation methods, time since death, preservation method (i.e. buffers vs. cold temperatures), and storage temperature of various buffers over time. Our results suggest that putting tissue into a buffer prior to freezing may be better than directly into ultra-cold conditions.

Greater than *X* kb: A quantitative assessment of preservation conditions on genomic DNA quality, and a proposed standard for genome-quality DNA.

Daniel G. Mulcahy, Kenneth S. Macdonald III, Seán G. Brady, Christopher Meyer, Katharine Barker, and Jonathan Coddington

Global Genome Initiative, National Museum of Natural History, Smithsonian Institution, Washington, DC USA

Corresponding Author:
Daniel Mulcahy
National Museum of Natural History, 1000 Constitution Ave., Washington, DC 20560
Email Address: MulcahyD@si.edu

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

Abstract

Advances in biodiversity genomic sequencing will increasingly depend on the availability of DNA samples—and their quantifiable metadata—preserved in large institutional biorepositories that are discoverable to the scientific community. Improvements in sequencing technology constantly provide longer reads, such that longer fragment length, higher molecular weight, and overall "genome-quality" DNA (gDNA) will be desirable. Ideally, biorepositories should publish numerical scale measurements of DNA quality useful to the user community. However, the most widely used technique to evaluate DNA quality, the classic agarose gel, has yet to be quantified. Here we propose a simple and economical method using open source image analysis software to make gDNA gel images quantifiable, and propose percentage of gDNA "greater than *X* kb" as a standard of comparison, where *X* is a band from any widely used DNA ladder with desirably large band sizes. We employ two metadata standards ("DNA Threshold" and "Percent above Threshold") introduced as part of the Global Genome Biodiversity Network (GGBN) Darwin Core extension. We illustrate the method using the traditionally used *Hin*dIII ladder and the 9,416 base-pair (bp) band as a standard. We also present data, for two taxa, a vertebrate (fish) and an invertebrate (crab), on how gDNA quality varies with seven tissue preservation methods, time since death, preservation method (i.e. buffers vs. cold temperatures), and storage temperature of various buffers over time. Our results suggest that putting tissue into a buffer prior to freezing may be better than directly into ultra-cold conditions.

Subjects Biodiversity, Bioinformatics, Genomics, Molecular Biology, Zoology

Keywords Agarose gels, DNA extractions, Genomic DNA, Tissue preservation

Introduction

Increasing the length of sequence reads is a core technological challenge in genomic science. Currently, the most widely used bench top technologies can achieve quality reads up to 150–500 bp in length (Glenn, 2011; Loman et al., 2012), but longer-read sequencing technologies are increasingly feasible (Loman and Quinlan, 2014). Therefore, access to high quality, high molecular weight DNA will become increasingly crucial. It is now feasible, practical, and increasingly more common to sequence complete genomes of non-model organisms (e.g. Pisani et al., 2015). As biodiversity genomics expands to rarer, harder-to-access, or vanishing organisms, obtaining "genomic quality" tissues—that provide high molecular weight DNA—becomes a significant challenge.

Natural history museums and academic institutions are currently obtaining, curating, and rapidly increasing biodiversity biobank collections (i.e. biorepositories), in order to maintain genomic quality material of non-model organisms, and to make this material available for scientific researchers conducting genomic analyses around the globe (Droege et al., 2014). In turn, making massive tissue and DNA collections discoverable is a priority for data aggregators, such as the Global Genome Biodiversity Network (GGBN; http://www.ggbn.org). GGBN is a network of institutions dedicated to preserving genetic resources, but also to advancing the data model for tissues, DNAs, RNAs, and similar resources and their standardizations (Droege et al., in press).

We present a simple, cost-effective agarose gel electrophoresis method for qualitatively analyzing genomic DNA (gDNA) extractions (for genomic analyses) that

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

can provide a quantifiable value of gDNA. Our method requires only basic molecular laboratory equipment (standard TBE gel rigs, UV imaging station, etc.); thus it does not rely on any expensive reagents or more expensive analytical quantification equipment (e.g. Spectrophotometers, Automated Electrophoresis Systems, Tape Stations, etc.).

Studies that explore the impact of tissue preservation on DNA quality often measure DNA concentration or test whether particular loci will PCR-amplify and sequence using traditional Sanger methods (e.g. Dawson et al., 1998; Vink et al., 2005; Yodder et al. 2006; Erkens et al., 2008; Frampton et al., 2008; Gaither et al., 2011; Moreau et al., 2013). However, these approaches are limited, as even fragmented DNA may amplify and produce high quality Sanger sequencing products, particularly when these products are in the size-range for Sanger methods [e.g. ~500–1000 base-pairs (bp)]. DNA concentration can be increased by adding more material (i.e. tissue) or combining multiple extractions from the same material, and does not provide any information about size. For genomic quality DNA, it is preferred to be mostly intact (whole chromosomes and organelle genomes), particularly if the intention is to sequence entire genomes, as the assembly of degraded gDNA (non-randomly) sheared gDNA prior to library preparation can be problematic for most NextGen platforms (Chen et al., 2015). Therefore, in order to assess the quality of their DNA, many researchers use agarose gels with high molecular-weight DNA ladders to visualize size and quality (Williams, 2007; Gaither et al., 2011), in addition to more sophisticated methods such as spectrophotometry, flourometry, or automated electrophoresis methods.

Here, we propose a standardized, simple method for electrophoresing genomic DNA on agarose gels with the standard λ phage *Hin*dIII ladder. The size of the gDNA

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

can be assessed by comparison to any of six bands in the *Hin*dIII ladder (2,027 bp, 2,322 bp, 4,361 bp, 6,557 bp, 9,416 bp, and 23,130 bp), and from this comparison, the percent of gDNA greater than a given band size can be calculated from a regular gel image. We recommend this method to standardize quality assessment of tissues collected and reporting by biorepositories and data aggregators such as GGBN. We suggest use of the ~"9 kb" (= 9,416 bp) size marker as a working standard, because it is substantially longer than standard NextGen reads (e.g., Illumina, etc.), and given current technologies, would be an appropriate minimum for long read sequencing. However, any size marker can be used to quantify gDNA, hence the method is referred to as "greater than *X* kb."

We present our method with a case study exploring preservative methods within field collection workflows that yield DNA of sufficient quality and quantity suitable for genomic sequencing. Within this case study, fresh, field collected tissue of a fish (*Morone americanus*) and a crab (*Callinectes sapidus*) were used to test if DNA quality is dependent on treatment and time until preservation. Alternative preservative solutions, temperature, and time were used as variables. Immediate cryopreservation was used as a benchmark for comparison as most researchers to date believe that freezing tissue at ultra-cold temperatures, such as -80°C or -190°C (liquid nitrogen), is the best preservative method for yielding genomic quality DNA.

We address four questions in this study: 1. Can DNA quality (in terms of fragment length) be measured quickly, consistently, and economically? 2. How does preservation method (buffers vs. temperature) affect DNA quality? 3. How does time since death affect DNA quality? 4. How does storage temperature (in various buffers) affect DNA quality?

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

Materials & Methods

Evaluating Preservation Methods 1.0

Tissue Collection 1.1

Muscle tissue was collected from two species: *Morone americanus* (white perch, hereafter "fish") and *Callinectes sapidus* (blue crab, hereafter "crab") at the Smithsonian Environmental Research Center (SERC) in Edgewater, MD. The crabs were euthanized using liquid nitrogen asphyxiation (held above liquid nitrogen, which depletes oxygen), and the fish were euthanized with MS222, following our ACUC protocols. Individual fish were filleted immediately after death. Fillets were cut into small strips (~0.5 cm x 5 cm) and immediately submerged in liquid nitrogen (LN_2). Individual crabs were dismembered upon death and claws were immediately submerged in LN_2. Two experiments were conducted ("Time" and "Temperature"; see below), and because this involved processing nearly 500 samples, all samples were submerged into LN_2 within five minutes after death to reduce postmortem tissue degradation. Each experiment only used the tissue collected from a single individual. For both experiments, each treatment combination consisted of ten replicates.

Time Experiment 1.2

Fillet strips and claws were thawed and sub-sampled; samples were weighed to the nearest mg before being subjected to one of seven preservation treatments: 1 ml 95%

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

EtOH, 1 ml salt-saturated DMSO-EDTA buffer (modified from Seutin et al., 1991), 1 ml DNAzol (Molecular Research Center), 1 ml of RNA*later* (Ambion), 300 μl M2 tissue digestion buffer (Autogen, Inc.), frozen at -20° C, frozen in LN_2 (≈-190°C). The DMSO-EDTA buffer we use is a slight modification of the tissue buffer used by Seutin et al. (1991); we use 25% of DMSO, instead of 20% (with 25% of 0.5M EDTA, 50% sterile H_20, saturated with sodium chloride). Before being subjected to the preservation treatment, all samples were allowed to sit at room temperature for one of three time periods: preserved immediately after thawing (<10 minutes total thaw time), 3 hours after thawing, and 24 hours after thawing. After 14 (crab) or 20 (fish) days, all sample tubes were moved into LN_2 and stored until DNA was extracted.

Temperature Experiment 1.3

Tissue samples were placed into one of five preservatives: 1 ml 95% EtOH, 1 ml DMSO-EDTA, 1 ml DNAzol, 1 ml RNA*later*, 300 μl M2 lysis buffer and then stored for 15 (crab) to 21 (fish) days at one of five temperatures: room temperature (RT), 4°C, -20° C, -80° C, and ~-190° C (LN_2). Subsequently, all sample tubes were stored in liquid nitrogen until DNA was extracted.

DNA Extraction 1.4

All samples were digested overnight in 300 μl AutoGen M2 and 300 μl M1 buffer (including Proteinase K). DNA was extracted from 300 μl (½ of the digested amount) of each digested sample by an AutoGen Prep 965 automated DNA extractor (AutoGen Inc.) using the manufacturers standard animal tissue (phenol-chloroform) extraction method,

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

and then dried. Samples were eluted in either 100 μl (fish) or 50 μl (crab) R9 DNA re-suspension solution (AutoGen Inc.).

DNA Quantification 1.5

All sample extractions were quantified through fluorescence, using a BioTek Synergy HT Multi-Mode Microplate Reader and Quant-iT dsDNA Assay Kit, broad range (Invitrogen, Cat# Q33130). Eight μl of eluted DNA was added to 200 μl of buffer and 1 μl of reagent in an opaque black 96-well microplate (Corning, Cat# 3915) and mixed thoroughly. Ten μl of seven solutions with known dsDNA concentrations (0, 10, 20, 40, 60, 80, and 100 ng/μl) were each added to two wells on every plate to calculate standard concentration curves. After sitting at room temperature for three minutes, samples were excited at 485 nm and ensuing fluorescence was read at 528 nm. Each sample was read twice, with five minutes between reads, and reads were averaged. Duplicate fluorescence values from concentration standards were averaged, and a general linear model of these fluorescence values versus total known DNA amounts was calculated using the program R, with the intercept constrained to run through the origin. This model was then used to calculate the total ng of DNA in each sample. The DNA concentration of the extraction was calculated by dividing total DNA by eight (the volume of sample used). We also calculated the total DNA extracted by multiplying fluorescent sample total DNA by 12.5 (for fish, because it was eluted in 100 μl) or 6.25 (for crab, because it was eluted in 50 μl). Finally, we calculated a DNA extraction yield (ng DNA/mg Tissue) by dividing the total DNA extracted by the weight of each tissue

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

sample, and multiplying by two (because only half of each digest was used in the extraction).

Statistical Analyses 1.6

Treatment differences were evaluated separately for each species (fish, crab) and experiment (Time, Temperature). The program R was used to run two-factor Analyses of Covariance (ANCOVA) on the quality and quantity datasets separately. Analyses included a time or temperature by preservative interaction term and used tissue weight as a covariate. If weight was non-significant, it was removed and the same ANCOVA, but without the covariate, was run. In total, 14 analyses were conducted for each experiment. Therefore, a Bonferroni correction of $\alpha = 0.004$ was used as a measure of significance.

Gel Quantification: Greater than X *kb 2.0*

Gel electrophoresis protocol 2.1

Extracted gDNA for all samples was visualized on a 1% agarose Tris-Borate-EDTA (TBE) gel. Five μl of each gDNA extract was loaded into the gel, and electrophoresed at 45 volts for 2.5 hours in 1X TBE buffer. To estimate gDNA fragment length, 0.5–1 ug of *Hin*dIII ladder was loaded into wells on each side of the DNA-loaded wells. After electrophoresis, gels were stained for 30 minutes in a solution of Ethidium Bromide (EtBr; at a final concentration of 0.5 μg/ml) and 1X TBE buffer. Gels were subsequently de-stained (to reduce background staining) in H_2O for 15 minutes. Finally, gDNA was visualized, photographed and images were stored as TIFF files using a

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

198 Syngene Gene GeneGenius Bio Imaging System.

199 Additionally, because EtBr is carcinogenic, and many labs are moving away from
200 its use and replacing it with safer methods, we also optimized the gel electrophoresis
201 protocol with GelRed™ (BioTium). For this method, we recommend a 0.7% agarose
202 Tris-Borate-EDTA (TBE) gel, run for ~2.5 hours at 45 volts in 1X TBE buffer. When
203 using GelRed™, we recommend loading a few wells of diluted *Hin*dIII ladder in various
204 amounts (e.g. 1:24, 1:49, and 1:99) because the GelRed™ can cause wide smearing in the
205 *Hin*dIII ladder bands (Fig. 1). For either staining method, it is important to have *Hin*dIII
206 ladder on either ends of the gDNA, such that it "brackets" the gDNA samples on either
207 end for post-scoring of the gel images (Fig. 1).

208

209 Scoring of the Gels 2.2

210 *Specific, step-by-step instructions can be found on the online Supplementary Information*
211 *Appendix I.*

212

213 To visualize and score the gel images, we used the program ImageJ v1.48 (W.
214 Rasband, NIH). ImageJ is an open-source, Java based program in the public domain,
215 available at http://rsb.info.nih.gov/ij. Gel image files are opened in ImageJ, with the
216 wells of the gel at the top. The gel image is color inverted to enhance visualization of the
217 bands and make density curves positive, rather than negative (see below). The image is
218 made level with respect to the top band in each ladder on either side of the gDNA
219 samples, the background is subtracted to remove smooth continuous backgrounds from
220 the gel images. Next, a vertical box is drawn encompassing the entire length of the lane

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

containing *Hin*dIII ladder, starting below the well, above the ~23 kb band, and to the extent of the gDNA (Fig. 1A). Vertically, the box should encompass the entire length of the smallest fragments of the gDNA lane on the gel with the greatest range.

The *Hin*dIII ladder box on the left is selected as the first lane. If using multiple dilutions of *Hin*dIII ladder, select the dilution that has the clearest bands on both sides of your gDNA samples. The box over the *Hin*dIII ladder is then moved to the first gDNA lane, and this lane is selected as the next to be analyzed (note: when moved, the original box stays in place and a new one is placed over the next lane). The horizontal location of this box must be carefully selected, so that the bands (or smears) of gDNA encompass the entire width of the box. The program automatically adjusts the vertical placement to be level with the first box. Then, the box from the first gDNA lane is moved to the next gDNA lane to be scored (typically, the next one to the right), again carefully selecting the location of the new lane horizontally, centering the box on the gDNA. Additional gDNA lanes are selected from left to right, by dragging the previous box from the left to the right, and selecting "Next Lane," until all desired lanes are included. The last lane selected must be the *Hin*dIII ladder to the right of gDNA lanes (Fig. 1A). An "Intensity Plot" is then created of the selected lanes, including the *Hin*dIII ladder lanes.

The Intensity Plot opens in a new window and is rotated 90° clockwise from the gel orientation (i.e. the leftmost lane becomes the top intensity plot). A straight line is drawn for the DNA size threshold (Fig. 1B) from the apex of the ladder threshold peak (e.g. "9 kb") on the first ladder (top of the Intensity Plot) to the apex of the ladder threshold peak on the second ladder (bottom of the Intensity Plot). This separates the intensity curves of each lane into a region greater than the threshold peak (Fig. 1C) and a

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

region less than the size of the threshold peak (Fig. 1D). All regions must be closed to be measurable; if the right side of the intensity curve does not meet the vertical line at the right side of the plot (leaving this region open), the Straight tool is used to draw a vertical line connecting the right part of the curve to the border of the plot.

The Wand tool is then used to select and measure the area of a region under the Intensity Plot, on the left side of the vertical line (>9 kb) and to the right of the vertical line (<9 kb). Once selected, the area is automatically calculated and presented in a Results table (Fig. 1E). If a dark imperfection appears in the gel that is clearly not part of the gDNA (Fig. 1A, gDNA II), a peak is recorded in the Intensity Plot. Similar to closing areas to measure (described above), one can eliminate the erroneous peak by using the Straight tool (Fig. 1F), and recalculate the area (Fig. 1G). The numbers in the Results window are the areas of the curve greater and less than the size of the threshold peak, respectively (Fig. 1E). These data are then copied and pasted into a data processing file and the percentage of the area greater than the chosen peak (e.g. ~ 9 kb) is easily calculated by dividing the area to the left (> 9 kb), by the total area.

Scoring Tests 2.3

We ran two analyses to test the repeatability (the variation obtained when one person measures samples repeatedly using the same methods) and reproducibility (the variation obtained when multiple people measure samples repeatedly using the same methods) of our gel scoring method. In the first analysis, two co-authors (CM and KSM) each independently scored the same gel image (FishTime <10 min) consisting of 40 lanes of gDNA, 10 times. For each scoring, the entire process was repeated, starting with

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

opening the raw image in ImageJ. Additionally, each scoring process was timed to give an estimate of method efficiency. The results of the 20 scored gel images (consisting of 800 scored gDNA lanes) were analyzed using a Gage repeatability and reproducibility ANOVA (Gage R&R) using the spreadsheet devised by J. Muelaner (*www.muelaner.com/quality-assurance/gage-r-and-r-excel/*). In order to test the consistency of dDNA in the gels, for the second analysis we ran seven gDNA samples, four crab (from Time <10 min) and three fish (from Temp = 4°C), each multiple times on two different gels. Each crab sample was run three times on each gel, while each fish sample was run four times on each gel. The gel images were also independently scored by two co-authors (DM and KSM), and the results of this test were also analyzed using a Gage R&R.

Results

Time Experiment

Figure 2A shows the fish gDNA extractions run out on a gel with the *Hin*dIII ladder from the seven preservation methods at time <10 mins after death. The DMSO-EDTA and DNAzol buffers have the greatest percent of gDNA > 9 kb (72% and 87%, respectively), consistent with the gel patterns showing the largest bands of gDNA, with little streaking or smearing in the lanes, indicating very little fragmented DNA. Figure 2B shows the results of the fish gDNA extractions from samples preserved at room temperature in 95% EtOH at <10 minutes, three hours, and 24 hours after death before being frozen in LN_2.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

The gDNA degrades through time resulting in little to no high molecular weight DNA after 24 hours (Fig. 2B).

Figure 3 shows the gel image of the crab gDNA for time <10 mins., for seven different preservation methods. The EtOH, DMSO, and DNAzol gDNAs have the greatest percentage of gDNA > 9kb (93%, 93%, 100%, respectively), and show less smearing and larger fragment size than RNAlater, -20°C, -190°C, and M2.

Average measures of quantities, concentrations, and quality (as measured by % of DNA > 9,416 bp) are shown for each trial for the time experiment for the fish and crab gDNA extractions in Table 1. Figures 4 and 5 show the quality of fish and crab gDNA (% > 9 kb), respectively, for seven different preservation methods over three time periods. DNA quality varies greatly in both taxa at time <10 mins., but all methods show degradation in quality of gDNA over time for both fish and crab tissues.

Temperature Experiment

Figure 6A shows the fish gDNA extractions electrophoresed on a gel with the *Hin*dIII ladder from five different preservation methods at room temperature. Note the DMSO-EDTA, DNAzol, and RNAlater buffers have the greatest percentage of gDNA > 9kb (89%, 98%, 100%, respectively), and show the largest bands of gDNA with little streaking or smearing in the lanes, indicating very little fragmented gDNA. Figure 6B shows the results of the fish gDNA extractions electrophoresed on a gel (with the *Hin*dIII ladder) from the five different temperatures the tissue were stored at while in the DMSO-

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

EDTA buffer. Figure 7 shows the gel image of the crab gDNA for the different preservation methods stored at room temperature. Average measures of quantities, concentrations, and quality (as measured by % of DNA > 9,416 bp) are shown for each trial for the temperature experiment for the fish and crab gDNA extractions in Table 2. Figures 8 and 9 show the quality of fish and crab gDNA (% > 9 kb), respectively, for the different preservation methods over the five temperatures at which tissues were stored.

ANCOVA Statistics

Table 3 shows the results of the analysis of covariance statistics for the fish and crab, Time and Temperature experiments, for both quality (% of gDNA > 9 kb) and quantity (ng gDNA/mg tissue) of gDNA. In the fish Time Experiment, time, preservation method, and the interaction variable all significantly affected DNA quality, while only time had a significant affect on quantity. For both quality and quantity, the covariate weight was marginally insignificant. In the crab Time Experiment, all factors (time, preservative, time x preservative, weight) significantly affected gDNA quality. The interactive term was significant only for gDNA quantity, although preservation method was just marginally insignificant. For the fish Temperature Experiment, only preservation method significantly affected gDNA quality, while temp, preservation method and weight all affected quantity. Finally, in the crab Temperature Experiment, only preservation method significantly affected either quality or quantity, although weight was marginally insignificant for both (Table 3).

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

Repeatability and Reproducibility

The first (single gel) test resulted in the repeatability variation (the % of total variation that is attributable to the scorer, i.e. the variability among the 10 scores a single co-author gave the same sample) of 3.99% and reproducibility variation (the % of total variation attributable to differences in the way the co-authors scored the same sample) of 5.71%. The total Gage Repeatability and Reproducibility variation was 6.97%. The second (multiple gel) test had a repeatability variation of 22%, a reproducibility variation of 18.1%, and a total Gage R&R variation of 28.5%.

Discussion

Greater than X kb

Here we demonstrate a simple, consistent, and efficient method for determining the size and quality of genomic DNA that does not require expensive equipment or reagents. Previous studies have presented the effects of different preservation conditions on DNA without providing an objective metric for genomic quality as we have done (e.g., Gaither et al., 2011; Camacho-Sanchez et al., 2013). We propose this method as a heuristic standard for biodiversity biobanking facilities and the genomic community, which may desire an initial assessment of DNA quality before requesting tissue samples. Using this method, genomic DNA can be electrophoresed on an agarose gel with a *Hin*dIII ladder,

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

or any other large-sized DNA marker, and quantified using simple plots in the free software ImageJ (W. Rasband, NIH: http://rsb.info.nih.gov/ij.).

We tested whether a single, or multiple researchers could reliably score a gel image similarly multiple times using a using a Gage Repeatability and Reproducibility ANOVA. The Gage R&R evaluates repeatability by the amount of variation attributable to a single measurer through multiple measures of the same sample, and evaluates reproducibility by the amount of variation attributable to differences between measurers. The analysis also calculates an overall measure of repeatability and reproducibility (the total Gage R&R), where most guidelines consider any total Gage R&R values under 10% to be acceptable, and any total values under 30% to be acceptable under certain conditions (Pan, 2006). For this test, the repeatability variation was 3.99%, the reproducibility variation was 5.7%, and the Total Gage R&R was 6.97%. Both researchers gave the same sample image similar quality scores (though not surprisingly, quality scores were slightly more different between researchers than were scores by the same researcher) to an extent acceptable to most quality control applications.

We conducted a second test to evaluate quality scores from samples run multiple times on a gel, and on separate gels. Our Gage R&R variability was much higher for this experiment, with a repeatability variation of 22%, a reproducibility of 18%, and a Total Gage R&R variation of 28.5%. There are many factors that can affect these scores in addition to measurer variability, such as variation in image quality (camera exposure, focus, dynamic range), variation in gel staining (length of time, mixing of stain, type of stain), and pipetting variation. Minor nuances between gel runs can result in slight

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

discrepancies of quality scores, but in our case most values of the same sample were within ~5% of individual scores.

For the purposes of demonstrating the method, we chose the 9,416 bp (~"9 kb") size marker from *Hin*dIII as the standard, and reported the percentage of genomic DNA greater than 9 kb, with the recognition of 50% or more of the gDNA being greater than 9 kb as a candidate indicator of "genomic quality." Although other fragment sizes could be chosen, in our experience of legacy biorepository samples, many gDNA extractions will fail to meet a higher standard. NextGen sequencing techniques are capable of much longer reads than 9 kb (Loman and Quinlan, 2014); therefore, threshold measures of genomic quality will be useful to the field of biodiversity genomics.

Moreover, this method does not depend on the choice of fragment size as a threshold for "genomic quality." One could as well pick the *Hin*dIII 564, 2,027 or other fragment sizes as a standard. From the point of view of a biodiversity tissue and DNA repository, whose samples may be been collected years ago and under difficult field conditions, or whose future samples may require difficult field conditions, we propose that 9 kb is, given current technology, a pragmatic value. Whatever the standard chosen, the threshold percentage of gDNA also implies that the extraction will contain fragments much larger than the actual threshold value. Importantly, the "DNA Threshold" and the "Percent above Threshold" standards (http://terms.tdwg.org/wiki/GGBN_Data_Standard - GGBN_Gel_Image_Vocabulary) provide a computable number for comparative values. These values coupled with the gel images allow the researcher to reach their own conclusions on the quality of gDNA for their specific needs.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

Of course, genomic DNA of many small organisms, such as certain arthropods, nematodes, meiofauna, and other microscopic organisms is generally difficult to visualize on agarose gels, yet suitable amounts of genomic sequence data can be successfully amplified from such organisms (e.g. Blaimer et al., 2015). We also realize that "degraded" gDNA, < 9 kb for example, can still be used for myriad analyses (e.g. sequence capture, ultra-conserved elements, etc.), including complete genome sequencing, such as the Neanderthal genome (Prüfer et al., 2014). Indeed, ancient DNA rarely exceeds 100 bp. However, as biodiversity scientists seek to preserve samples from all major clades of the tree of life, from all biomes, practical and economical field techniques must be developed, and in turn, the effectiveness of such techniques should be quantitative.

Whole genome sequencing will advance technically to use very long fragment sizes, as longer reads provide higher quality assemblies (Schatz et al., 2010). Therefore, for plants and animals that can easily be visualized on an agarose gel, we recommend the "*greater than X kb*" method as a standard for biodiversity biobanking laboratories to report the quality of gDNA extracts.

Typically, most library preparation methods to date begin with shearing gDNA to sizes compatible with the maximum size range of most NextGen sequencing platforms (e.g. 300–500 bp). Therefore, one might question why we should be concerned with large pieces of intact gDNA prior to library preparation. Mechanical shearing, or sonication, shear gDNA randomly across the genome, whereas degradation can cause shearing in non-random places, and in the same places repeatedly, possibly leading to biased NextGen results (Zackin and Ge 2010; Choi et al., 2002). Furthermore, the use of

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

large insert mate-pair libraries up to 25 kb can increase the efficiency of genomic structure analyses (van Heesch et al., 2013).

Preservation Methods

In our tests, salt-saturated DMSO/EDTA buffer and DNAzol are better at preserving high-quality (> 9 kb) gDNA than other methods such as direct storage in liquid nitrogen (-190° C) or -20° C storage (Fig. 2A and Fig. 3; Table 3). Saturating tissues with storage buffer immediately is also important, as significant DNA degradation can occur, even within three hours time after death (Fig. 2B and Figs. 4–5; Table 3). Temperature appears to have less of an effect on tissue preservation for overall size-quality of gDNA (Figs. 8–9; Table 3). Therefore, time before preservation and preservation method (buffer vs. frozen), and interactions between these factors, have the biggest influence on gDNA quality when measured as size, for both the fish and crab tissue samples (Table 3).

Currently, many genetic researchers working on non-model organisms are under the impression that directly freezing fresh tissue is the best way to preserve gDNA, and the faster and colder the method of preservation, the better (e.g. Wong et al., 2012). Liquid nitrogen can be expensive, and both liquid nitrogen and dry ice can be difficult to obtain and transport in certain countries and under remote field conditions. Our results show that putting tissue directly into buffers, such as the salt-saturated DMSO/EDTA or DNAzol is actually better than directly into liquid nitrogen or -20° C storage, without any buffer for fish (Figs. 4; Table 1), and putting tissues directly into buffers or liquid

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

449 nitrogen alone is far better than -20° C storage for the crab (Fig. 5); albeit all of our
450 samples were flash-frozen first in liquid nitrogen without any buffers prior to treatment.
451 This is good news for molecular biologists collecting field samples of genomic
452 material. Salt-saturated DMSO/EDTA is easy and inexpensive to make in the lab and is
453 more easily transported than reagents such as 95% EtOH. We suspect, based on our
454 observations, that the best method of preservation is to allow fresh tissue material to soak
455 in the salt-saturated DMSO/EDTA buffer for approximately 1 hr. (depending on amount
456 of tissue) and then preserving it in liquid nitrogen, or -20° C for transportation and/or
457 long term storage.
458 Interestingly, 95% EtOH appears to be just as good as DMSO for preserving
459 DNA quality in crab tissue (Fig. 5), but not so for fish (Fig. 4). Ethanol can cause
460 extensive, crude dehydration of animal tissues, which may cause fragmentation of gDNA
461 (Gaither et al., 2011). Some research supports that ethanol performs better in invertebrate
462 tissues (Williams, 2007), perhaps in insects because it can more easily penetrate the
463 cuticle and exoskeletons. Our crab tissue was removed from the shell prior to
464 preservation, which may have improved the performance of EtOH in our study.
465 Regardless, the DMSO/EDTA buffer and 95% EtOH each performed better than direct
466 cryopreservation for the fish and crab tissues, respectively.
467 Challenges of sufficient amounts of total DNA extracted and concentration (ng
468 DNA/mg tissue) can be overcome by increasing the amount of starting material, and/or
469 combining extractions from several separate extractions of the same starting material
470 source. Here, we have shown that time since death is the biggest factor in gDNA
471 concentration for the fish, but this does not seem to be a factor for the crab tissue (Table

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

3). Temperature is important for the preservation of fish tissue in terms of quantity (Table 3).

Our study was limited to a small number of organisms, one vertebrate, the white perch (*Morone americana*) and one marine invertebrate, the blue crab (*Callinectes sapidus*). Ideally, we would like to see our methods tested for a variety of organisms, including vascular plants, algae, and terrestrial invertebrates. Ultimately, we would like to see tests of methods of genomic DNA preservation for all groups of life.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

Acknowledgements

All or portions of the laboratory and/or computer work were conducted in and with the support of the L.A.B. facilities of the National Museum of Natural History (NMNH) or its partner labs. Specifically, we thank A. Driskell and N. Agudelo for help with designing and implementing the DNA extraction tests, and L. Weigt, J. Hunt, and M. Kweskin for L.A.B. facilities and resources. We thank C. Baldwin (NMNH) and R. Aguilar (SERC) for help with obtaining specimens, S. Whittaker for use of highly sensitive balance, and V. Gonzales for comments on the manuscript.

Supplemental Information

Supplemental information for this article can be found online at *(to be included upon publication; at the end of this pdf for review)*.

References

Blaimer BB, Brady SG, Schultz TR, Lloyd JW, Fisher BL, and Ward PS. 2015. Phylogenomic methods outperform traditional multi-locus approaches in resolving deep evolutionary history: a case study of formicine ants. BMC Evolutionary Biology (2015) 15:271. DOI: 10.1186/s12862-015-0552-5.

Camacho-Sanchez M, Burraco P, Gomez-Mestre I, Leonard JA. 2013. Preservation of RNA and DNA from mammal samples under field conditions. Molecular Ecology Resources 13: 663–673.

Chen TW, Gan RC, Chang YF, Liao WC, Wu TH, Lee CC, Huang PJ, Lee CY, Chen YYM, Chiu CH, Tang P. 2015. Is the whole greater than the sum of its parts? De novo assembly strategies for bacterial genomes based on paired-end sequencing. BMC Genomics (2015) 16:648. DOI: 10.1186/s12864-015-1859-8.

Choi, HJ, Lim ST, Lai PY, Chan CK. 2002. Turbulent Drag Reduction and Degradation of DNA. Physical Review Letters 89, 088302. DOI: http://dx.doi.org/10.1103/PhysRevLett.89.088302.

Dawson MN, Raskoff KA, Jacobs DK. 1998. Field preservation of marine invertebrate tissue for DNA analyses. Molecular Marine Biology and Biotechnology 7:145–152.

Droege G, Barker K, Seberg O, Coddington J, Benson EE, Berendsohn WG, Bunk B, Butler C, Cawsey EM, Deck J, Döring M, Flemons P, Gemeinholzer B, Güntsch A, Hollowell T, Kelbert P, Kostadinov I, Kottmann R, Lawlor RT, Lyal C, Mackenzie-Dodds J, Meyer C, Mulcahy DG, Nussbeck SY, Tuama ÉÓ, Orrell T, Petersen G, Robertson T, Söhngen C, Whitacre J, Wieczorek J, Yilmaz P, Zetzsche H, Zhang Y, Zhou X. The Global Genome Biodiversity Network (GGBN) Data Standard specification. *In press*. Database.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

Erkens, RHJ, Cross H, Maas JW, Hoenselaar K, Chatrou LW. 2008. Assessment of age and greenness of herbarium specimens as predictors for successful extraction and amplification of DNA. Blumea 53:407–428. DOI: 10.3767/000651908X608052.

Gaither MR, Szabó Z, Crepeau MW, Bird CE, Toonen RJ. 2011. Preservation of corals in salt-saturated DMSO buffer is superior to ethanol for PCR experiments. *Coral Reefs* 30:329–333. DOI 10.1007/s00338-010-0687-1.

Glenn TC. 2011. Field guide to next-generation DNA sequencers. Molecular Ecology Resources (2011) 11:759–769. DOI: 10.1111/j.1755-0998.2011.03024.x.

Frampton M, Droege S, Conrad T, Prager S, Richards MH. 2008. Evaluation of specimen preservatives for DNA analysis of bees. Journal of Hymenoptera Research 17:195–200.

van Heesch S, Kloosterman WP, Lansu N, Ruzius FP, Levandowsky E, Lee CC, Zhou S, Goldstein S, Schwartz DC, Harkins TT, Guryev V, Cuppen E. 2013. Improving mammalian genome scaffolding using large insert mate-pair next-generation sequencing. BMC Genomics 14:257. DOI: 10.1186/1471-2164-14-257.

Kilpatrick CW. 2002. Non-cryogenic preservation of mammalian tissues for DNA extraction: an assessment of storage methods. Biochemical Genetics 40:53–62.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

546 Loman, NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, and Pallen
547 MJ. 2012. Performance comparison of benchtop high-throughput
548 sequencing platforms. Nature Biotechnology 30:434–439. DOI: 10.1038/nbt.2198.
549
550 Loman NJ, Quinlan AR. 2014. Poretools: a toolkit for analyzing nanopore sequence
551 data. Bioinformatics 30:3399–3401. DOI: 10.1093/bioinformatics/btu555.
552
553 Moreau CS, Wray BD, Czekanski-Moir JE, Rubin BE. 2013. DNA preservation: a test
554 of commonly used preservatives for insects. Invertebrate Systematics 27:81–86. DOI:
555 org/10.1071/IS12067.
556
557 Pan, J-N. 2006. Evaluating the Gauge Repeatability and Reproducibility for Different
558 Industries. Quality & Quantity 40:499–518. DOI: 0.1007/s11135-005-1100-y.
559
560 Pisani D, Pett W, Dohrmann M, Feuda R, Rota-Stabelli O, Philippe H, Lartillot N,
561 Wörheide G. 2015. Genomic data do not support comb jellies as the sister group to all
562 other animals. Proceedings of the National Academy of Sciences. 112:15402–15407.
563 DOI: 10.1073/pnas.1518127112.
564
565 Prüfer, K, Racimo F, Patterson N, Jay F, Sankararaman S, Sawyer S, Heinze A, Renaud
566 G, Sudmant PH, de Filippo C, Li H, Mallick S, Dannemann M, Fu O, Kircher M,
567 Kuhlwilm M, Lachmann M, Meyer M, Ongyerth M, Siebauer M, Theunert C, Tandon A,
568 Moorjani P, Pickrell J, Mullikin JC, Vohr SH, Green RE, Hellmann I, Johnson PFL,

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

569 Blanche H, Cann H, Kitzman JO, Shendure J, Eichler EE, Lein ES, Bakken TE,
570 Golovanova LV, Doronichev VB, Shunkov MV, Derevianko AP, Viola B, Slatkin M,
571 Reich D, Kelso J, Pääbo S. The complete genome sequence of a Neanderthal from the
572 Altai Mountains. *Nature* 505:43–49. DOI: 10.1038/nature12886.
573
574 Schatz, MC, Delcher AL, Salzberg SL. 2010. Assembly of large genomes using second-
575 generation sequencing. Genome Research 20:1165–1173. DOI: 10.1101/gr.101360.109.
576
577 Saavedra-Matiz CA, Isabelle JT, Biski CK, Duva SJ, Sweeney ML, Parker AL, Young
578 AJ, DiAntonio LL, Krein LM, Nichols MJ, Caggana M. 2013. Cost-effective and
579 scalable DNA extraction method from dried blood spots. Clinical Chemistry 59:1045–
580 1051. DOI: 10.1373/clinchem.2012.198945.
581
582 Seutin, G, White BN, Boag PT. 1991. Preservation of avian blood and tissue samples for
583 DNA analyses. Canadian Journal of Zoology-Revue Canadienne de Zoologie 69: 82–90.
584
585 Vink CJ, Thomas SM, Paquin P, Hasashi CY, Hedin M. 2005. The effects of
586 preservatives and temperatures on arachnid DNA. Invertebrate Systematics 19:99–104.
587 DOI: org/10.1071/IS04039.
588
589 Williams ST. 2007. Safe and legal shipment of tissue samples: does it affect DNA
590 quality? Journal of Molluscan Studies 73:416–418. DOI: 10.1093/mollus/eym039.
591

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

592 Wong PBY, Wiley EO, Johnson WE, Ryder OA, O'Brien SJ, Haussler D, Koepfli K-P,
593 Houck ML, Perelman P, Mastromonaco G, Bentley AC, Venkatesh B, Zhang Y, Murphy
594 RW. 2012. Tissue sampling and standards for vertebrate genomics. GigaScience 1:8.
595 DOI: 10.1186/2047-217X-1-8.
596
597 Yodder M, Tandingan de Ley I, King IW, Mundo-Ocampo M, Mann J, Blaxter M, Poiras
598 L, De Ley P. 2006. DESS: a versatile solution for preserving morphology and
599 extractable DNA of nematodes. Nematology 8:367–376.
600
601 Zackin JL, Ge W. 2010. Polymer and surfactant drag reduction in turbulent flows. In:
602 Utracki LA, Jamieson AM. eds. *Polymer Physics: From Suspensions to Nanocomposites*
603 *and Beyond*. Hoboken NJ, John Wiley & Sons Inc., 89–128.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

Tables

Table 1. Time Experiment. Average quantities of Extracted DNA (ng), DNA extraction yield (ng DNA/mg tissue digested; see text for method of calculation) and Quality (% DNA > 9,416 bp) of genomic DNA extracted from tissues of two species of fish (*Morone americana*) and crab (*Callinectes sapidus*) held in one of 21 treatments: three times prior to preservation (<10 min, 3 hrs, 24 hrs), then stored in seven preservative methods (95% EtOH, DMSO-EDTA, DNAzol, RNAlater, M2, -20°C, and -190°C) for a minimum of 14 days. Mean and standard deviation (SD) values are shown.

	< 10 min						3 hrs						24 hrs					
	Extracted DNA (ng)		Extraction Yield (ng DNA/mg Tissue)		Quality (% > 9416 bp)		Extracted DNA (ng)		Extraction Yield (ng DNA/mg Tissue)		Quality (% > 9416 bp)		Extracted DNA (ng)		Extraction Yield (ng DNA/mg Tissue)		Quality (% > 9416 bp)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
	Fish: *Morone americana*																	
EtOH	186	28	78	10	6	3	209	49	83	18	4	4	98	14	40	7	1	2
DMSO	183	57	81	20	72	10	171	94	71	40	25	8	13	5	6	2	0	0
DNAzol	244	170	108	71	87	5	157	25	63	12	34	8	14	4	6	2	0	0
RNAlater	182	17	81	8	7	6	180	36	72	20	0	0	86	37	38	18	0	0
M2	214	44	101	21	35	20	174	31	79	13	4	4	84	18	40	8	3	5
-20°C	236	37	102	15	7	5	166	29	69	11	1	2	110	45	49	19	2	5
-190°C	191	44	84	15	3	2	165	24	70	10	0	2	137	43	59	18	0	0
	Crab: *Callinectes sapidus*																	
EtOH	97	48	22	10	93	5	369	312	79	64	46	36	303	94	61	17	1	1
DMSO	38	11	8	3	93	13	136	85	30	15	48	22	28	15	5	2	0	0
DNAzol	34	28	8	7	100	0	130	108	28	23	20	10	14	4	3	1	0	0
RNAlater	140	74	28	14	81	17	231	126	52	30	53	37	184	33	35	7	0	0
M2	119	88	34	26	75	18	250	114	64	29	56	32	31	40	8	10	0	0
-20°C	275	56	58	12	17	16	278	143	70	43	26	37	196	66	43	7	0	0
-190°C	209	112	44	23	72	16	381	214	90	47	49	36	102	141	19	26	0	0

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

Table 2. Temperature Experiment. Average quantities of Extracted DNA (ng), DNA extraction yield (ng DNA/mg tissue digested, see text for method of calculation) and Quality (% DNA > 9416 bp) of genomic DNA extracted from tissues of two species fish (*Morone americana*) and crab (*Callinectes sapidus*) held in one of 25 treatments: five storage temperatures (Room Temperature = RT, 4°C, -20°C, -80°C, and -190°C) x five preservative methods (95% EtOH, DMSO-EDTA, DNAzol, RNAlater, M2) for a minimum of 14 days. Mean and standard deviation (SD) values are shown.

	RT						4°C						-20°C					
	Extracted DNA (ng)		Extraction Yield (ng DNA/mg Tissue)		Quality (% > 9416 bp)		Extracted DNA (ng)		Extraction Yield (ng DNA/mg Tissue)		Quality (% > 9416 bp)		Extracted DNA (ng)		Extraction Yield (ng DNA/mg Tissue)		Quality (% > 9416 bp)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Fish: *Morone americana*																		
EtOH	233	129	150	69	46	18	229	87	157	54	23	12	188	38	120	23	41	11
DMSO	375	170	248	96	89	10	295	116	196	66	89	8	309	155	208	108	94	6
DNAzol	298	177	194	109	98	4	297	114	192	74	83	15	294	143	202	90	96	6
RNAlater	173	127	125	102	100	0	169	44	109	23	59	14	160	28	111	13	81	6
M2	161	73	110	52	82	4	180	33	134	18	30	17	188	87	142	64	80	10
Crab: *Callinectes sapidus*																		
EtOH	1238	354	430	112	52	35	889	190	329	81	62	24	804	432	291	159	39	14
DMSO	888	367	320	155	79	6	388	162	141	54	37	20	727	345	257	98	45	27
DNAzol	643	252	242	100	90	10	705	436	257	177	77	27	511	371	171	110	NA	NA
RNAlater	641	175	229	56	62	23	704	242	271	108	55	11	399	131	142	50	54	7
M2	582	166	222	67	84	15	606	282	227	105	73	21	750	194	268	81	59	19

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

Table 2 (continued).

	-80°C						-190°C					
	Extracted DNA (ng)		Extraction Yield (ng DNA/mg Tissue)		Quality (% > 9416 bp)		Extracted DNA (ng)		Extraction Yield (ng DNA/mg Tissue)		Quality (% > 9416 bp)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Fish: *Morone americana*												
EtOH	159	31	105	17	49	24	306	105	209	71	32	17
DMSO	372	192	263	132	95	13	488	244	326	189	82	20
DNAzol	447	184	307	127	100	0	609	238	403	157	40	42
RNAlater	184	78	125	57	77	10	174	61	117	35	71	23
M2	164	45	116	31	88	6	269	66	172	33	63	13
Crab: *Callinectes sapidus*												
EtOH	1264	580	415	173	24	4	942	222	320	87	77	24
DMSO	864	472	308	186	24	3	715	284	241	101	49	28
DNAzol	920	270	332	110	28	3	432	213	151	72	91	12
RNAlater	772	667	289	271	36	16	532	320	178	90	65	24
M2	848	474	309	171	78	23	1024	273	336	93	60	28

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

Table 3. Summary of ANCOVA results for Time and Temperature experiments. Covariance was tested for between Time, Preservation method (Pres.), and Weight (Wt.) for the Time experiments and Temperature (Temp.), Preservation method (Pres.), and Weight (Wt.) for the Temperature experiments. Interactions (Int.) were tested for between Time and Temp. and Pres., if weight was not significant it was removed. A Bonferroni correction of $\alpha = 0.004$ is used. Significant results are shown in boldface font.

		Time Experiment				Temperature Experiment			
Fish:		Time	Pres.	Wt.	Int.	Temp.	Pres.	Wt.	Int.
Morone americana	Quality	**<0.001**	**<0.001**	0.0074	**<0.001**	0.0408	**<0.001**	0.2079	0.8923
		<0.001	**<0.001**		**<0.001**	0.0339	**<0.001**		0.9260
	Quantity	**<0.001**	0.0452	0.0743	0.0220	**<0.001**	**0.0026**	**<0.001**	0.7179
		<0.001	0.0505		0.043				
Crab:									
Callinectes sapidus	Quality	**<0.001**	**0.0022**	**0.0010**	**0.0014**	0.0197	**<0.001**	0.0607	0.5840
						0.0081	**<0.001**		0.5550
	Quantity	0.6902	0.0099	0.5916	<0.00191	0.7545	**<0.001**	0.0129	0.0568
		0.7591	0.0135		**0.0015**	0.4273	**<0.001**		0.0481

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

Figures

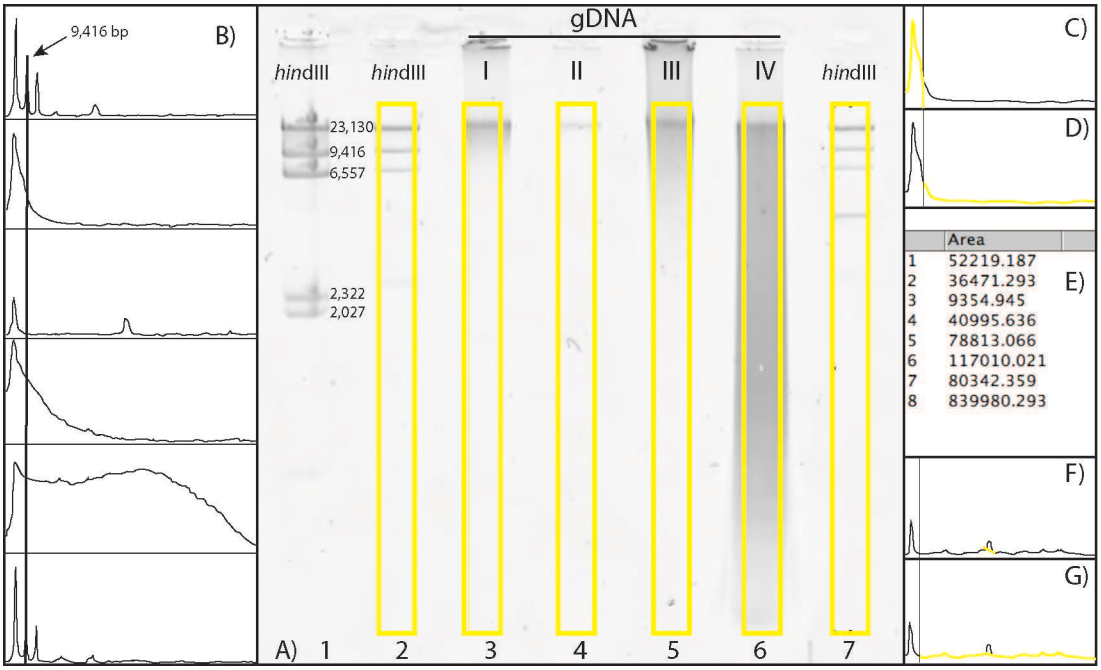
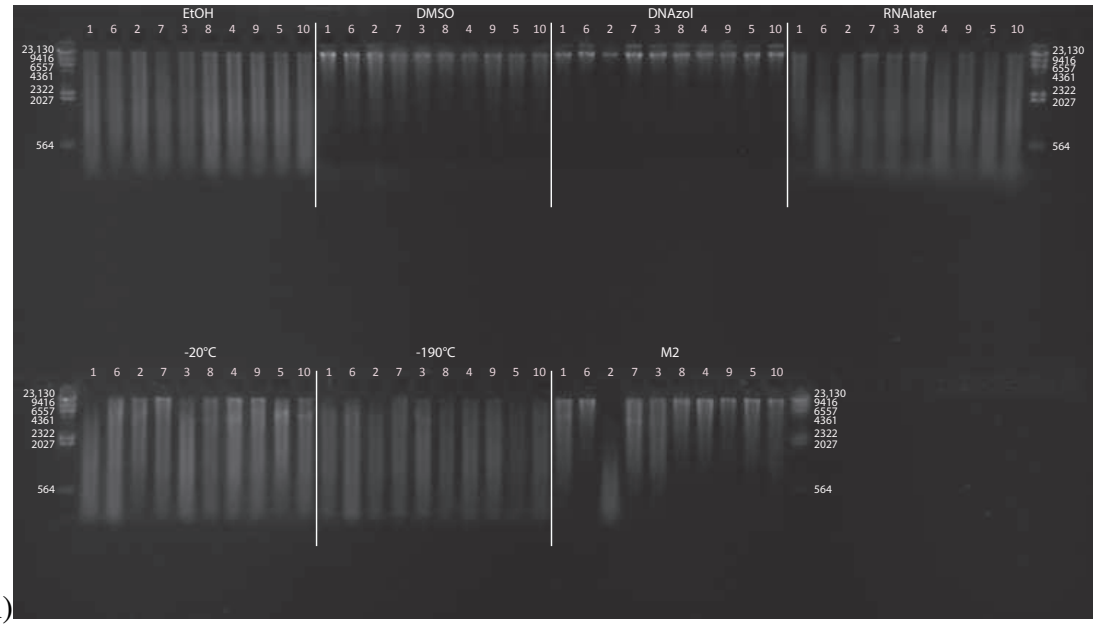


Figure 1. Scoring a gDNA gel in ImageJ. A) An inverted gDNA gel image with the *Hin*dIII ladder in wells 1–2 (wells labeled on bottom) and 7, and gDNA samples (I–IV) in wells 3–6. The bands in the *Hin*dIII ladder are labeled in well 1; note the ~2 kb bands are visible in this lane, but not in the more diluted lanes (2 and 7; see text). B) The Plot results of the six yellow boxes scored in A), with the Straight tool line spanning the ~9 kb peaks of the two *Hin*dIII ladders (wells 2 and 7, top and bottom, respectively), with the four gDNA samples in the middle. C) The area calculated for > 9 kb. D) The area < 9kb. E) The Results box, with values 1 and 2 for the areas > 9 kb and < 9kb for gDNA sample I, respectively. F) The Straight tool line used to eliminate peak cause by flaw in gel for gDNA sample II. G) The calculated area < 9kb with the erroneous peak removed. The percent of gDNA > 9kb is calculated by dividing the area > 9kb (C) by the total area below the curve (C + D).

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

A)



B)

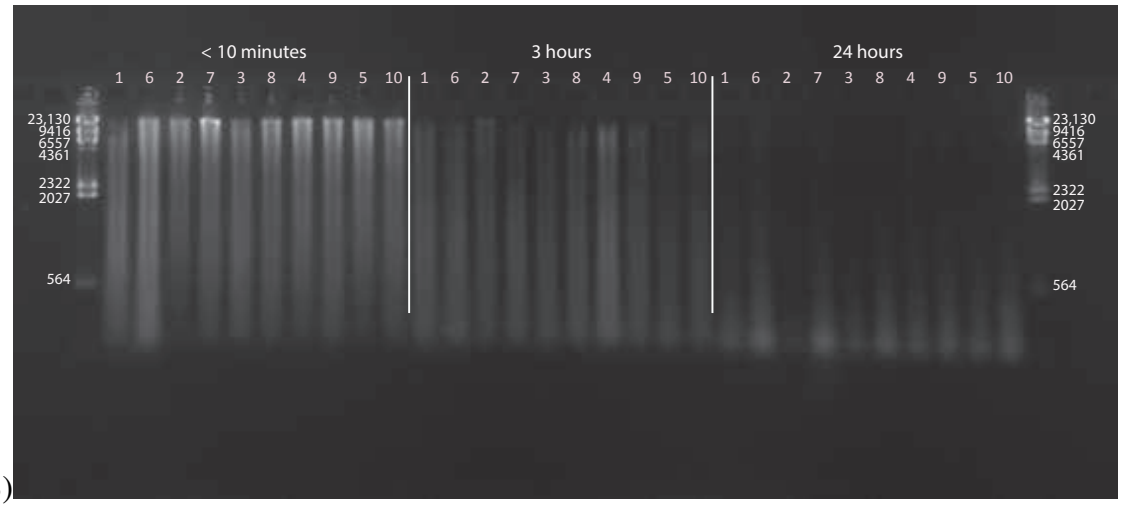


Figure 2. Gel image of extracted gDNA from the white perch, *Morone americana* for the Time Experiment. The ladder in the left- and right-most lanes is the *Hin*dIII with bands at 564, 2027, 2322, 4361, 6557, 9416, and 23130 bp. A) Samples in all tissue storage treatments for Time-since-death <10 minutes: 1) in EtOH; 2) in salt saturated DMSO/EDTA preservation buffer ("DMSO"); 3) submerged in DNAzol Reagent (Invitrogen); 4) submerged in RNAlater (Ambion); 5) submerged in M2 tissue digestion solution (Autogen); 6) held at -20°C with no preservation solution; 7) submerged in liquid nitrogen (≈-190°C) with no preservation solution. B) Samples in Time-since-death treatments for EtOH tissue storage treatment are shown for the three different time periods (< 10 mins., 3 hrs., 24 hrs.).

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

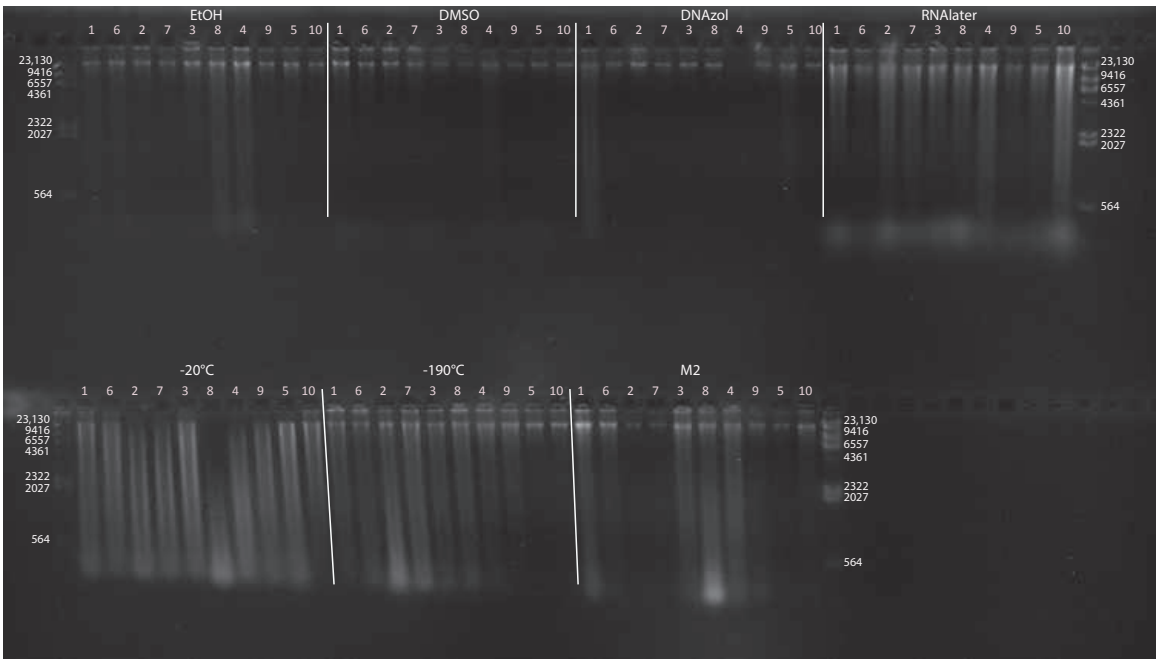


Figure 3. Gel image of extracted genomic DNA from the blue crab, *Callinectes sapidus*, for the Time Experiment, showing all tissue storage treatments for Time-since-death <10 minutes. The *Hin*dIII ladder is shown in the left- and right-most lanes of the gel.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

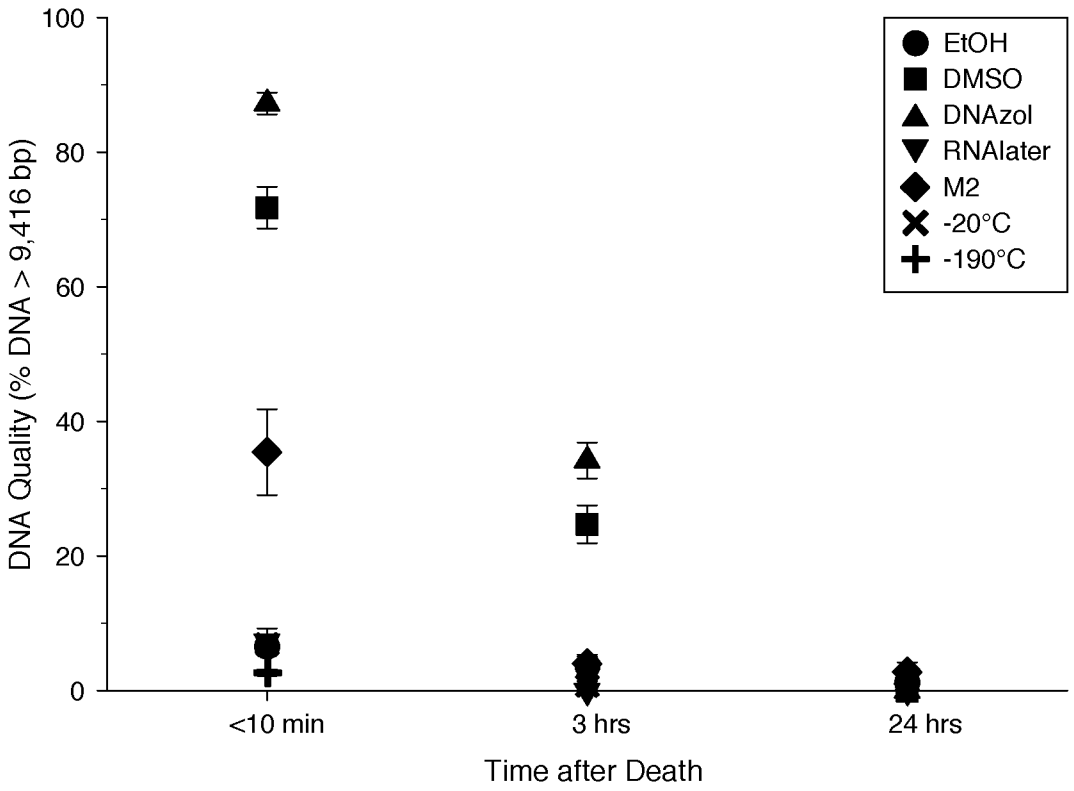


Figure 4. Graph of Fish Time Experiment. Quality of gDNA extracted from the white perch, *Morone americana*. Quality is shown as the mean % of gDNA > 9,416 bp on the y-axis. Preservation treatments are differentiated by symbols (see lgend). Time-since-death treatments are shown on the x-axis for the three time periods tissue samples sat before preservation. See Table 1 for exact values of each method.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

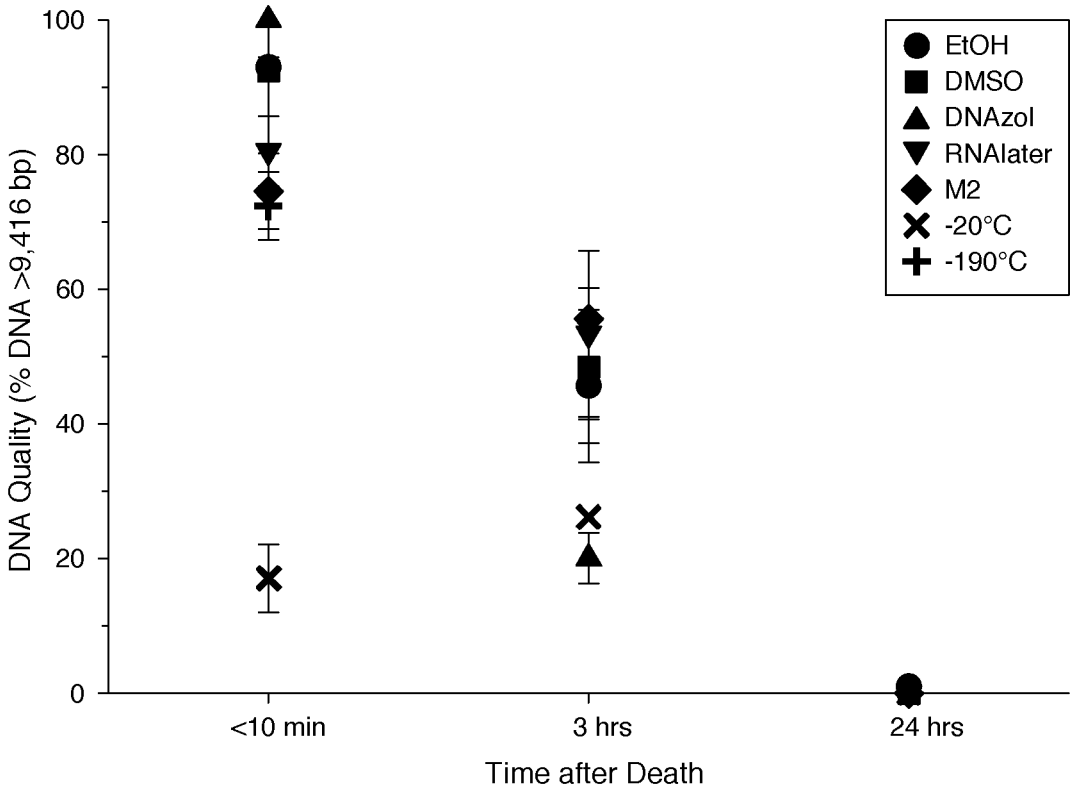
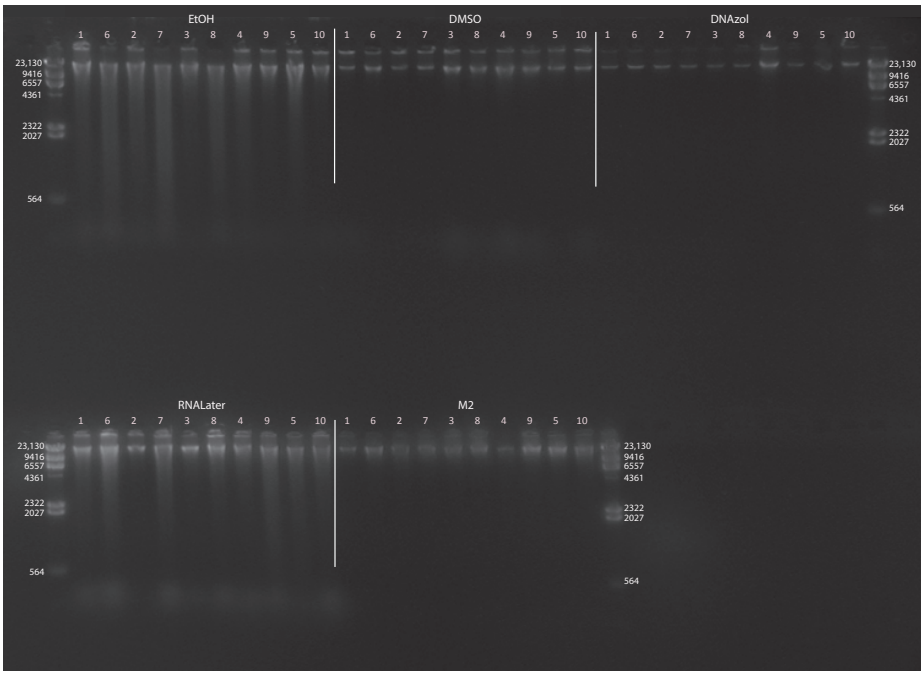


Figure 5. Graph of Crab Time Experiment. Quality of gDNA extracted from the blue crab, *Callinectes sapidus*. Quality of extracted gDNA is shown as the mean % of DNA > 9,416 bp on the y-axis. Preservation treatments are differentiated by symbols. Time-since-death differentiated are shown on the x-axis for the three time periods tissue samples sat before preservation. See Table 1 for exact values of each method.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

A)



B)

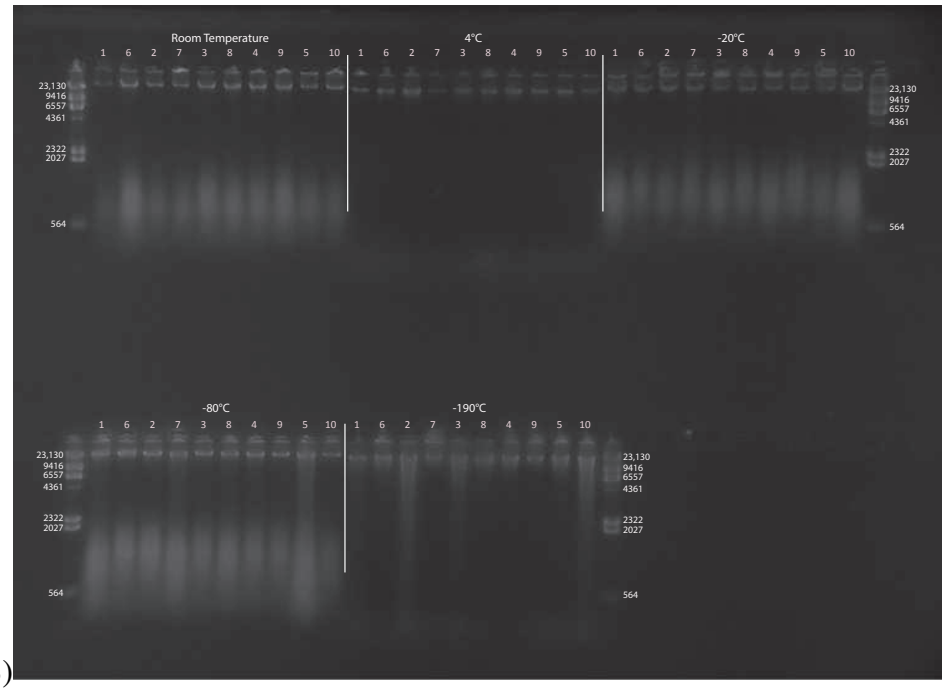


Figure 6. Gel image of extracted gDNA for the white perch, *Morone americana*, for the Temperature Experiment. Prior to DNA extraction, tissue was stored for 14–20 days in one of five solutions (EtOH, DMSO, DNAzol, RNAlater, M2) and kept at one of five temperatures: Room Temperature, -20°C, -80°C, -190°C. A). Showing all tissue storage buffer treatments for tissue storage at Room Temperature. B). Showing all tissue storage temperature treatments for DMSO-EDTA salt buffer.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

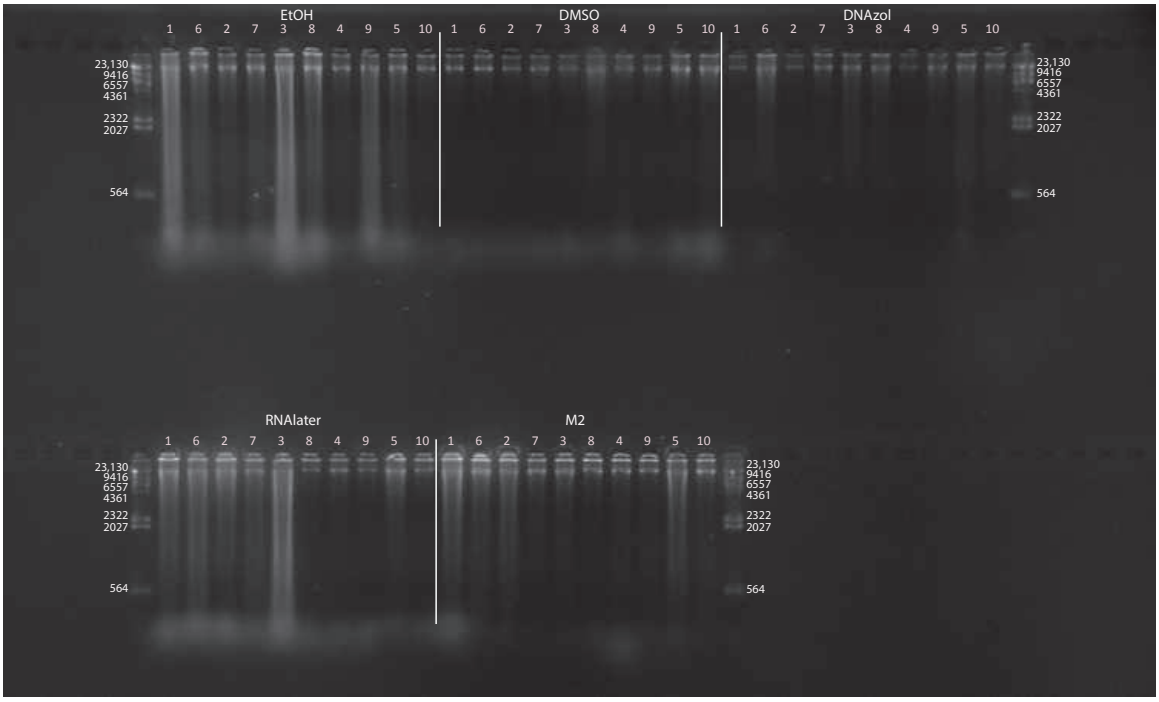


Figure 7. Gel image of extracted genomic DNA for the blue crab, *Callinectes sapidus* for preservation Temperature Experiment, showing all tissue storage buffer treatments for tissue storage at room temperature.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

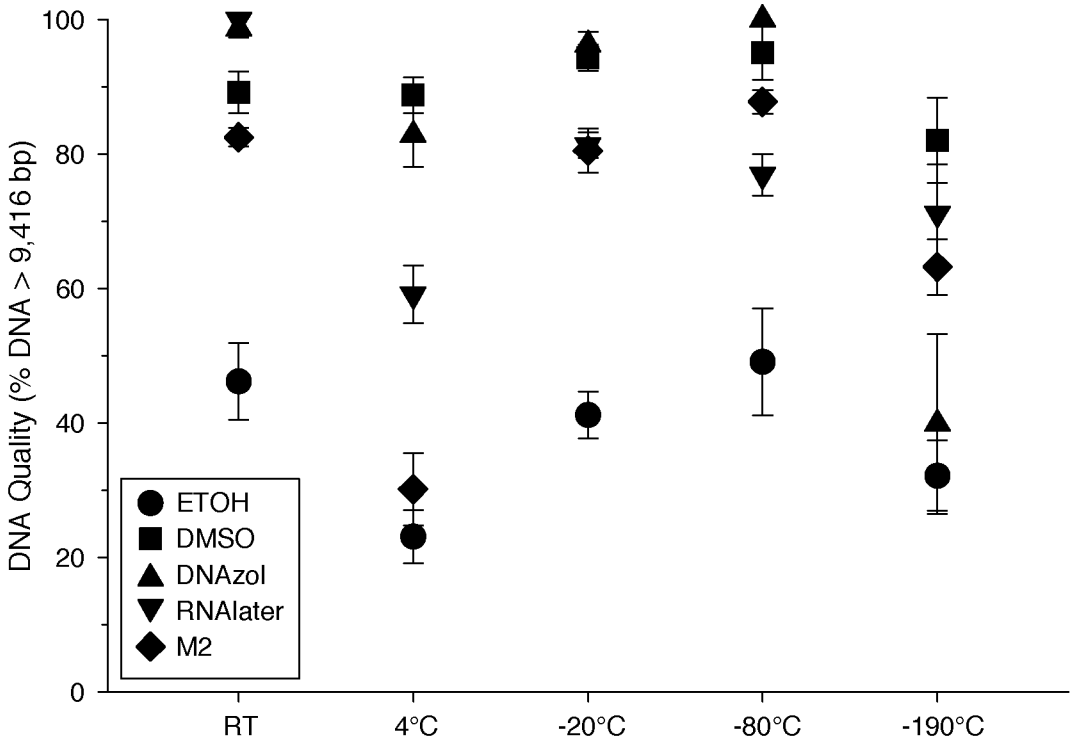


Figure 8. Graph of Fish Temperature Experiment. Quality of gDNA extracted from white perch, *Morone americana*. Quality of extracted gDNA is shown as the mean % of gDNA > 9,416 bp on y-axis. Preservation solutions are differentiated by symbols (see legend), with the different preservation temperatures on the x-axis. See Table 2 for exact values of each temperature.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

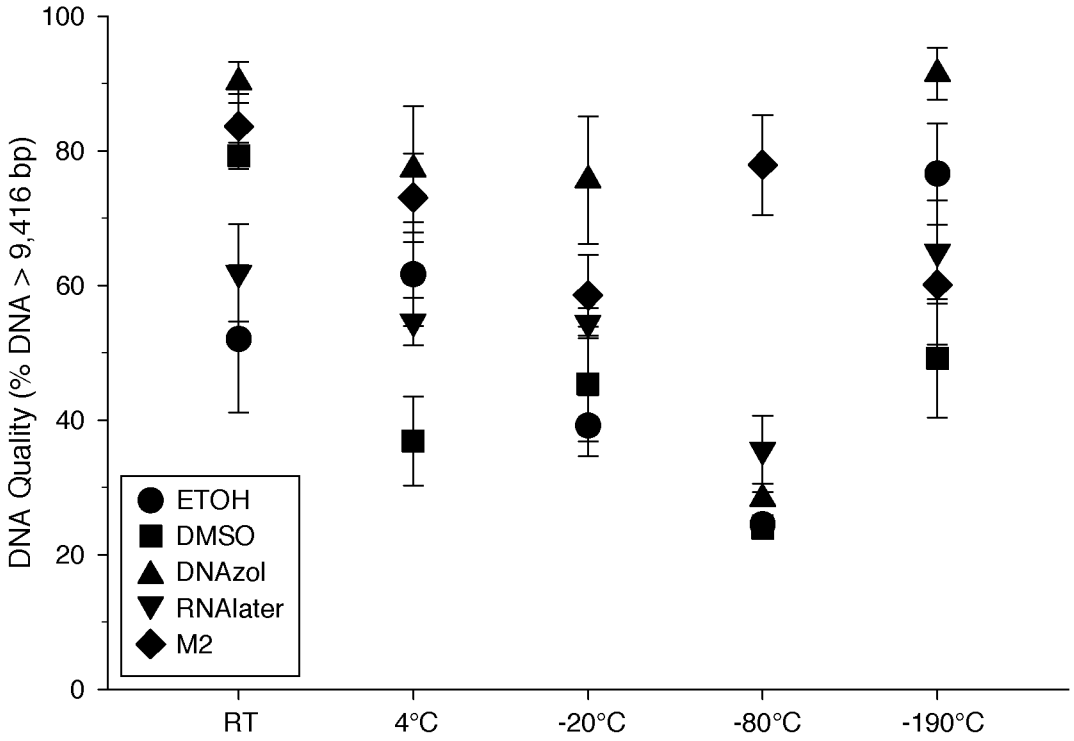


Figure 9. Graph of Crab Temperature Experiment. Quality of gDNA extracted from the blue crab, *Callinectes sapidus*. Quality of extracted gDNA is shown as the mean % of gDNA > 9,416 bp on the y-axis. Preservation treatments are differentiated by symbols. Temperatures are shown on the x-axis for the five temperatures tissue samples were held after preservation. See Table 2 for exact values of each temperature.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2202v1 | CC BY 4.0 Open Access | rec: 1 Jul 2016, publ: 1 Jul 2016

Supplementary Information: Mulcahy et al., PeerJ

Protocol for using ImageJ to determine relative band size of DNA from gel images

1. **Download Program**: Download Java program ImageJ from http://rsb.info.nih.gov/ij
2. **Open Image**: Open raw image file in ImageJ with ⌘o or using the menu (File>Open).
3. **Invert Image**: Invert the image with shift+i. This is not necessary, but I find that it is easier to see the limits of bands with an inverted image, and it makes the density curves that we will create positive rather than negative. It also makes Step 4 (image leveling) significantly easier.
4. **Level Image**: Make sure lanes in your image are horizontally level using the menu (Image>Transform>Rotate). This brings up a dialog that allows you to rotate the image by degrees (positive numbers for clockwise rotation, negative numbers for counterclockwise rotation). To determine if rotation is necessary, select the Preview box, set the Angle to 0, and increase the Grid Lines value. This allows you to compare your lanes to the horizontal and vertical lines of the grid. Change the Angle value until the lanes in your image are horizontally level (i.e. you want the top bands in both ladders to be level).
5. **Subtract Background**: Subtract background from image using the menu (Process>Subtract Background). A pop-up window will open. Make sure that the box for Light Background is checked, as well as the box for Preview. Choose the Rolling ball radius that gives the whitest background without degrading DNA signal. Usually the default (50.0 pixels) works fine, but you may want to play with this. If you use too low of a value, you start to lose dark pixels at the edges of the DNA bands, but you want the lowest value obtainable without losing dark pixels.
6. **Select Rectangle Tool**: Select the Rectangle tool from the toolbar
7. **Draw Box Encompassing Ladder**: Draw a rectangle on the lane of the first ladder. The bands of the ladder should cross the entire width of the drawn box, so make the box as narrow as possible. Vertically, the box should encompass the entire DNA band of the lane on your gel with the greatest vertical range (but do not include the well in the box). In other words, there should be no white space on the left or right sides between the DNA band and the sides of the drawn box, but there should be white space between the longest band and the top of the box, and white space between the shortest band and the bottom of the box. If you make a mistake in drawing boxes or selecting lanes and need to start over, this can be done via the menu (Analyze>Gels>Reset)
8. **Select Box as First Lane**: Select this lane as the first lane to be analyzed using ⌘1. This can also be accomplished using the menu (Analyze>Gels>Select First Lane)
9. **Select Next Lane**: Drag the box to the next lane to be measured, and select this lane as the next to be analyzed using ⌘2. This can also be accomplished using the menu (Analyze>Gels>Select Second Lane). When you drag the box to the

new lane, carefully select the location of the new lane horizontally, attempting to center the box on the bands (as in Step 8). You do not have to be as careful with vertical placement of the box: ImageJ will automatically align new boxes vertically as you select them.

10. **Select all Lanes:** Repeat Step 9 for all lanes that you wish to analyze. Make sure that the last lane selected is the last ladder. It doesn't matter which ladder is selected first, and which is selected last, as long as they are the first and last selected.
11. **Create Intensity Plot**: Create intensity plot of the selected lanes using ⌘3. This can also be accomplished using the menu (Analyze>Gels>Plot). If you need to remake the intensity plot, close the current plot, make the gel image the active window, and use the menu (Analyze>Gels>Re-Plot Lanes). This will give you a new, clean intensity plot.
12. **Draw Line for DNA Size Cutoff:** From the toolbar, choose the Straight tool. Draw vertical line on the intensity plot from the apex of the ladder peak of choice on the first ladder (top of page) to the apex of the ladder peak of choice on the second ladder (bottom of page). This separates the intensity curves of each lane into a region greater than the size of the peak chosen and a region less than the size of the peak chosen.
13. **Close All Regions to Measure**: All regions under the intensity curves must be closed to measure area, but you will notice that the right side of the intensity curve does not meet the vertical line at the right side of the plot, leaving this region open. To fix this, with the Vertical tool still selected, draw a vertical line from the furthest right part of the curve on the first ladder to the furthest right part of the curve on the last ladder. This may have to be repeated on the left side of the curve if it doesn't meet the left side of the plot. You will know if this is necessary if, during Step 14, the Wand tool highlights more than just the area under the curve.
14. **Select Region Above Cut-off for First Measurement**: Choose the Wand tool from the toolbar. Select the region under the intensity curve on the left side of the vertical line from Step 12 for your first sample.
15. **Measure First Region**: Some users may find that an area measurement is automatically taken when the region is selected using the Wand tool. If a new box (labeled "Results") appears when you select a region, then measurements are automatic. If no "Results" box appears, you need to tell ImageJ to measure each region with ⌘m, or using the menu (Analyze>Measure). The number in the results window is the area of the curve greater than the size of the peak chosen. If you need to measure multiple areas (i.e. the curve reaches the lower limit of the graph, essentially splitting up the region), select each while holding shift, then measure after all regions have been selected. If measurements are automatic, you cannot select multiple regions before measuring. Instead, measure each region separately and add them together to get the area of the entire curve
16. **Select and Measure Region Below Cut-Off**: Repeat Steps 14 and 15 with the region under the intensity curve on the right side of the vertical line. This is the area of the curve less than the size of the peak chosen, and should appear in the "Results" window, below the first value. Note that ImageJ numbers the

measurements sequentially as they are made, so keep track of which value belongs to each curve area measurement.

17. **Measure All Regions**: Repeat Step 14 through Step 16 for each sample.
18. **Save Results**: Save the Results table with ⌘s or using the menu (File>Save As). By default, ImageJ saves the table as a tab-delimited file with an excel (.xls) extension.
19. **Calculate Proportion:** For each sample, add the two portions (left and right) to get the total area under the curve. Divide the first area value (area to the left of the peak) by the total area to obtain the % of area under the curve greater than the size chosen.

Notes:
This protocol is for usage on a Mac, usage on a PC is similar (replace ⌘ with Ctrl for keyboard shortcuts).

The ImageJ users guide can be viewed at http://rsb.info.nih.gov/ij/docs/guide/index.html or downloaded using http://rsb.info.nih.gov/ij/docs/guide/user-guide.pdf.
This protocol was based on the video tutorial found at http://imagejdocu.tudor.lu/doku.php?id=video:analysis:gel_quantification_analysis
All video tutorials can be accessed at http://imagejdocu.tudor.lu/doku.php?id=video:start