Vortex fluidics-mediated DNA rescue from formalin-fixed museum specimens

Christian A. Totoiu[1,#a], Jessica M. Phillips[4], Aspen T. Reese[5], Sudipta Majumdar[1], Peter R. Girguis[5], Colin L. Raston[4], Gregory A. Weiss[1,2,3 *]

[1] Department of Chemistry, University of California, Irvine, California, United States of America

[2] Department of Molecular Biology and Biochemistry, University of California, Irvine, California, United States of America

[3] Department of Pharmaceutical Sciences, University of California, Irvine, California, United States of America

[4] Flinders Institute for Nanoscale Science and Technology, College of Science and Engineering, Flinders University, Adelaide, South Australia, Australia

[5] Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts, United States of America

[#a]Department of Chemical Engineering and Biotechnology, University of Cambridge, Cambridge, England, United Kingdom

[*] Corresponding Author:

E-mail: gweiss@uci.edu (GAW)

Abstract

DNA from formalin-preserved tissue could unlock a vast repository of genetic information stored in museums worldwide. However, formaldehyde crosslinks proteins and DNA, and prevents ready amplification and DNA sequencing. Formaldehyde acylation also fragments the DNA. Treatment with proteinase K proteolyzes crosslinked proteins to rescue the DNA, though the process is quite slow. To reduce processing time and improve rescue efficiency, we applied the mechanical energy of a vortex fluidic device (VFD) to drive the catalytic activity of proteinase K and recover DNA from American lobster tissue (*Homarus americanus*) fixed in 3.7% formalin for >1-year. A scan of VFD rotational speeds identified the optimal rotational speed for recovery of PCR-amplifiable DNA and while 500+ base pairs were sequenced, shorter read lengths were more consistently obtained. This VFD-based method also effectively recovered DNA from formalin-preserved samples. The results provide a roadmap for exploring DNA from millions of historical and even extinct species.

Introduction

Archived biological samples offer an important source of genetic information for diverse fields including evolutionary biology, ecology, phylogenetics, biodiversity, and epidemiology [1-2]. Samples, from hydrated tissues to whole organisms, have historically been preserved in aqueous formaldehyde (3.7 to 4% solution of formaldehyde in water, termed formalin). In many cases, these specimens are the only remaining samples that could provide genetic information about the organisms, including their microbiomes, environments, diets, and other attributes – all from the moment of sample preservation [3-5]. This preservative, however, hinders DNA amplification and sequencing with the sample [6]. Thus, new methods to recover DNA from formalin-fixed specimens could advance our ability to access the genetic information in these samples, and advance our understanding of how organisms and ecosystems have responded to natural and anthropogenic changes over time. For example, formalin-fixed specimens in natural history museums could be used to elucidate the impact of environmental changes on the DNA of biological populations [1-2, 7]. DNA sequencing of such samples could address longitudinal, biological questions that may be impractical to address without the genetic information for the preserved specimens [2, 8-9].

For >150 years, formalin fixation has been used to effectively preserve hydrated specimens [7]. A vast repository of formalin-fixed samples exists, including at least 400 million samples at 13 large institutions [1]. Marine organisms are particularly well-preserved in this aqueous preservative, as it retains morphological features well, enabling more detailed taxonomic studies. Aqueous formaldehyde is also advantageous in that it stops parasitic microbial growth [10]. However, preserving samples in formalin fixation damages DNA [11-12]. Covalent modification of DNA bases by the electrophilic formaldehyde drives base deglycosylation, and the resultant abasic sites in DNA can cause strand breakage [13]. Additionally, long duration storage often incurs DNA fragmentation, independent of formalin [14]. Fragments from both mechanisms increase the amount of DNA template required for PCR amplification of longer targets, and can also inhibit PCR [15-16].

Intrastrand and protein-DNA crosslinks introduced by formaldehyde can also block PCR and DNA sequencing [17-18]. Protein-DNA crosslinks result from nucleophilic attack on formaldehyde by proteins' primary amines to yield imines and iminium ions. These groups can then react with the less nucleophilic primary amines of DNA bases, particularly from guanine, resulting in a protein-DNA crosslink (Fig 1) [19-20]. Due to the high density of amines found on the surface of proteins and DNA, each DNA-protein complex can become crosslinked multiple times. Additionally, formalin-fixed cells cannot repair the slow process of cytosine deamination to uracil [11, 21]. During PCR amplification, adenine can be incorporated as the incorrect complement to degraded cytosine, resulting in point mutations [11]. In summary, DNA damage caused by preservation results in short templates for PCR and low-quality, error-prone DNA sequences.

Fig 1. Schematic of formalin-induced crosslink formation and the removal of crosslinked proteins by treatment with proteinase K. (A) A protein amine can nucleophilically attack the formaldehyde carbonyl to yield an iminium ion, which can then react with another primary amine from DNA, RNA, or proteins to form a crosslink. This crosslink reaction is in reversible dynamic equilibrium [10, 21]. **(B)** Treatment with a protease, proteinase K, allows free DNA (fDNA) recovery. Here, Nuc designates an amine nucleophile from the DNA.

Despite the immense challenges that are associated with formalin fixation, much effort has been dedicated to developing techniques for sequencing these irreplaceable samples (DNA recovery methods from formalin-fixed tissue are summarized in Table 1). Most current methods for recovering DNA from formalin-fixed organisms use proteinase K, a thermostable serine protease with broad substrate specificity [22], to digest crosslinked proteins and eliminate most crosslink-associated blockages [21-23]. However, even at the enzyme's optimum temperature (49 ± 2 °C), the free DNA (fDNA) recovery rate from this method is low at approximately 4.4% per hour; additionally, the enzyme's half-life is limiting at approximately 11.3 h [21-22]. At room temperature (≈22 °C), the proteolytic reaction rate yields only 1.1% fDNA per hour [24]. DNA can also be recovered from formalin-fixed tissue with a 0.1 M NaOH (pH 12) buffer treatment at 120 °C for 25 min [7, 25]. However, these harsh conditions can further damage the DNA through Brønsted base-caused strand cleavage; therefore, this approach is most valuable in cases where there is an excess of tissue to be digested, and is highly inappropriate for most delicate, longer-preserved samples. Notably, the current reactions to liberate DNA are harsh, low-throughput, low yielding, and time consuming. Mild methods to increase DNA recovery and purification for PCR amplification and subsequent DNA sequencing could revolutionize the study of a wide range of museum specimens.

Table 1. Current methods for formaldehyde crosslink removal & DNA recovery from formalin-fixed specimens.

Method	Temperature (°C)	Time (h)
proteinase K treatment[26]	56	~17
proteinase K treatment in Tris-NaCl-EDTA-SDS buffer[27]	55	68
hot alkali buffer treatment[7][25]	100 to 120	0.4 to 0.7
Cetyltrimethylammonium bromide (CTAB) & proteinase K[28]	65 & 56	0.5 & 1-72
QIAamp DNA Mini Kit[28]	56	not reported
QIAamp DNA FFPE Kit[28]	56 & 90	not reported

We posit that judicious application of mechanical energy could address this challenge. Specifically a vortex fluidic device (VFD) directs controlled mechanical energy into solution to accelerate enzyme-catalyzed reactions [29]. This thin film microfluidic platforms can disrupt membranes and drive protein folding [30], and potentially assist with the deaggregation, and solubilization of formalin-fixed samples in addition to acceleration of enzyme activity. Here, we tested the efficacy of using a VFD to

accelerate proteinase K activity, and increase the process throughput and efficiency of extracting DNA from formalin-fixed specimens (Fig 2). Thus, we optimized the recovery of fDNA from formalin-fixed specimens through VFD and post-processing purification. Our results suggest that this method recovers fDNA with 40 to 85% greater yields than conventional methods without requiring harsh conditions, and can decrease treatment time from days to hours.

Protocol

Reagents

- Proteinase K (Promega, cat. no. V3021, lyophilized)
- Sodium dodecyl sulphate (SDS) (Fisher Bioreagents, cat no. BP8200-5)
- Tris hydrochloride (Fisher Bioreagents, cat. no. BP153-1)
- Calcium chloride (Fisher Chemical, cat no. C614-500)
- Glycerol (ACS reagent, cat. no. G7893-4L)
- Ethylenediaminetetraacetic acid (EDTA) (Acros Organics, cat. no. 147850010)
- Nitrogen (Liquid) (Airgas Healthcare, cat no. UNI977)
- Ethanol (200 proof, Molecular Biology Grade) (Fisher Scientific, cat. no. BP2828-500)
- Clean & Concentrator Kit (Zymo Research, cat. no. D4006)

Equipment

- Vortex Fluidic Device v.2 (VFD) (Vortex Fluidic Technologies)
- Microliter pipettes (1000 µL, 200 µL, 20 µL, and 10 µL)
- Refrigerated, tabletop centrifuge
- Vortex mixer
- Mortar & pestle
- Hemostat
- Forceps
- Razorblades
- Eppendorf tubes (1.7 mL)

Reagent setup

Critical Step: Nanopure water (ddH_2O) is used for all buffers and solutions. Buffers are autoclaved or sterile-filtered, if containing SDS, prior to addition of enzymes and use. Enzyme-containing solutions are stored at -80 °C.

- **Proteolysis buffer** 20 mM Tris-HCl, 50 mM EDTA, 1% w/v SDS, pH 8.0[31]
- **Proteinase K solution** 10 mg/mL proteinase K, 20 mM Tris-HCl, 1 mM $CaCl_2$, 50% glycerol, pH 8.0

Procedure

Tissue sample preparation. Timing: ≈1 h

1. After removal of a small portion (≈5 to 10 g) of the preserved biological tissue from the preservative, immerse the tissue in liquid nitrogen until frozen, and grind with a mortar and pestle for ≈1 min. The small pieces of ground tissue are aliquoted (≈1 to 1.5 g) into Eppendorf tubes (henceforth termed tubes).

   Safety Note: Exercise caution when utilizing a sharp edge to prevent puncturing or cutting personal protective equipment or skin.

   Pause Point: At this stage, the sample can be stored at -80 °C for later processing.

2. On an autoclaved glass surface, mince the formalin-fixed tissue sample with a flame-sterilized razorblade held by a sterilized hemostat for ≈5 min.

   Safety Note: Exercise caution when utilizing a sharp edge to prevent puncturing or cutting personal protective equipment or skin.

   Critical Step: This step increases the surface area to volume ratio of the tissue and, thus, improves the proteinase K access to the sample.

3. To remove the preservative fluid, wash the sample three times with the proteolysis buffer (1 mL). For each wash step, briefly vortex, centrifuge (15 krcf, 3 min), and decant the samples. If required, an addition centrifugation (15 krcf, 1 min) can remove any residual buffer.

VFD treatment. Timing: ≈1 h

4. Transfer the ground, minced tissue to the bottom of an autoclaved 20 mm VFD sample tube. Add proteolysis buffer (950 µL) and then proteinase K solution (50 µL).
5. Seal the VFD sample tube with a rubber septum, and use the VFD to spin the sample (7 krpm, 1 h, RT).

   Critical Step: This rotational speed is optimal for fDNA amplification and sequencing.

Sample post-processing and purification. Timing: ≈1.5 h

6. Following VFD processing, immediately transfer the sample, including both the processed tissue and solution, from the VFD sample tube to a new tube.
7. Immediately, centrifuge the sample (15 krcf, 5 min, RT) to remove the tissue. Transfer the supernatant to a clean tube.
8. Incubate the supernatant on wet ice for 30 minutes.

   Critical Step: A white precipitate (SDS) collects at the bottom of the tube. The presence of SDS negatively affects PCR yields and subsequent purification of fDNA.
9. Immediately, centrifuge the sample (15 krcf, 10 min, 4 °C) to remove SDS, and transfer the supernatant to a new tube without disturbing the SDS pellet.

   Critical Step: The supernatant must be transferred immediately following centrifugation to prevent resolubilization of SDS.

   Pause Point: At this stage, the sample may be frozen at -20 or -80 °C for later analysis.
10. Process 400 µL of the supernatant with a Zymo DNA Clean & Concentrator Kit, according to the manufacturer's instructions.

fDNA quantification and characterization. Timing: ≈8 h

11. The isolated fDNA can be used for further amplification, characterization, quantification, and sequencing.

Materials & methods

One adult male Lobster (*H. americanus*) was purchased in February 2017 from a local lobster fishery in Boston, MA. The lobster was euthanized by quickly severing the ganglia behind the eyes with a sharp knife. The body was then placed whole in a solution of 3.7% formaldehyde in 0.9 M phosphate-buffered saline (which approximates the salinity of seawater). The lobster was maintained at room temperature for one month, and then shipped to the University of California Irvine in March 2017. All lobster trails shown here have used muscle recovered from the chelipeds (primary claws), which have remained in formalin for the two-year duration of this study. For experimental treatments, the lobster claw tissue was processed according to the procedure described above.

Proteinase K (Promega, V3021, lyophilized) was solubilized and diluted to 10 mg/mL in storage buffer (20 mM Tris-HCl, 1 mM $CaCl_2$, 50% glycerol, pH 8.0).

For DNA isolation, 100 (±1) mg of lobster tissue samples were utilized. Tissue preparation, VFD processing, and DNA purification used methods described above. In the experiments reported here, the VFD was operated in the confined, not continuous flow, mode [32-33]; specifically, 1 mL volumes were used. The negative control samples consisted of identical tissue samples and treatment but were not subjected to VFD processing. A positive control contained VFD-processed fresh lobster tissue. Additionally, fixed and fresh lobster tissue were processed overnight at 56 °C without the VFD for the conventional method controls, based on Table 1. Finally, intermediate controls combined the conventional and VFD-mediated methods to minimize the number of variables changed per experiment. These intermediate controls consisted of fixed and fresh lobster processed overnight (as in the conventional method) at room temperature (as in the VFD-mediated method) without use of the VFD.

The positive controls for PCR quantification were DNA obtained from fresh (non-formalin-fixed), ground lobster claw tissue. The DNA recovery from this sample applied Chelex 100 Resin (Bio-Rad, 10% w/v in 500 µL in ddH_2O) with the manufacturer's protocol. The mixture of a tissue fragment and resin was vortexed and centrifuged briefly before incubation at 90 to 95 °C for 20 to 35 min. Following another brief vortexing and centrifugation, the supernatant was isolated as the positive control for lobster fDNA.

DNA extraction yields were compared by quantitative PCR (qPCR) (Bio-Rad iCycler). For PCR, reaction mixtures (10 µL) applied the Phusion DNA polymerase (0.2 U final concentration, New England Biolabs) and buffer (5× diluted final concentration, New England Biolabs), DMSO (10% v/v final concentration), dNTPs (0.5 mM each final concentration, New England Biolabs), primers (8-33 ng each final concentration, Integrated DNA Technologies) (Table 2), and SYBR Green I dye (10,000× diluted final concentration, Thermo Fisher Scientific). A PCR was performed with 1 cycle of 94 °C for 5 min followed by 40 to 50 cycles of 94 °C for 1 min, 60 °C for 1 min, and 72 °C for 2 min, followed by 1 cycle of 72 °C for 5 min.

Table 2. PCR primer sequences and annealing temperatures.

Gene Target: mitochondrial ATP synthase		Primer	Annealing Temperature (est.) (°C)
579 bp	Forward	GGGTTACTTTTTATTCCCTACCTTTATTGAGC	60
	Reverse	GGCATATAAAGTCCTTAGAACAGCAAATACATACG	
183 bp	Forward	GGGTTACTTTTTATTCCCTACCTTTATTGAGC	60
	Reverse	CAGCCCGAGAGTGTTATTGAATATAATAAATC	

The recovered fDNA from the formalin-fixed lobster was quantified by UV-Vis absorbance. The absorbance spectra of the samples, diluted in ddH_2O (1:50), were measured (Jasco V-730 Spectrophotometer). The spectra were recorded from 200 to 400 nm, in triplicate (technical replicates), with a scanning speed of 200 nm/min and intervals of 0.5 nm with ddH_2O as the blank. Two buffer only controls examined the reaction mixture without lobster tissue. The first was processed in the VFD (7.5 krpm, 1 h, RT), and the second did not use VFD processing. The 1 kb Plus DNA Ladder (10 µg, 1000 µL) (Thermo Fisher Scientific, 1.0 µg/µl) was used to estimate DNA sizes and concentrations; a positive control applied DNA from fresh lobster tissue extracted with Chelex 100 Resin, as described above.

A SYBR Green I fluorescence assay also quantified the dsDNA concentration [34]. To derive a DNA concentration calibration curve, a 1 kb Plus DNA Ladder (Thermo Fisher Scientific, 1.0 µg/µL) was diluted to concentrations of 0 ng/µL to 1.25 ng/µL. These dilutions (100 µL) and SYBR Green I dye (100 µL, Thermo Fisher Scientific, 1250× diluted in ddH_2O) were added to a black, clear-bottom 96-well plate (Corning, 3615). The fluorescence of each well was measured (λ_{ex} 485 nm, λ_{ex} 550 nm, Bandwidth 20 nm, Gain 50). The fluorescence of fDNA from the fixed lobster samples (100× diluted) were similarly measured, and the dsDNA concentrations were estimated using the calibration curve shown in S4 Fig.

Results and discussion

Conventional methods applying proteinase K to remove protein-DNA crosslinks require >17 h [26-27], and typically the reaction runs for >1 day. Here we report enzyme acceleration techniques via VFD mechanical stimulation to enable DNA recovery in <2 h from lobster claw tissue (*H. americanus*) preserved in formalin (3.7% formaldehyde in phosphate-buffered saline, a condition iso-osmotic with seawater). Key variables requiring optimization included mechanical breakage of the tissue sample, length of fDNA sequences to be amplified, and rotational speed of the VFD. As observed for other VFD-enhanced enzymes [29], proteinase K activity can be accelerated using a VFD. The approach allows recovery of free DNA (fDNA) from formalin-fixed samples. The rotational speed of the tube in the VFD determines the level of shear, micro-mixing, and fluid dynamics experienced by the thin film of liquid [32]. The tilt angle of the VFD can be important, and 45° relative to the horizontal is the optimal tilt angle for a myriad of applications, including accelerating enzymatic reactions [29, 32].

Fig 2. Schematic of the VFD-mediated fDNA recovery technique. (A) The protocol begins with Vortex Fluidic Device (VFD) treatment (7 krpm, room temperature, abbreviated RT, 1 h) of a mixture of proteinase K and the frozen, then broken-up tissue. The reaction mixture is next processed to remove solids and DNA polymerase inhibitors. The recovered fDNA is then purified and concentrated. Finally, the DNA is amplified, quantified, and characterized by **(B)** qPCR and **(C)** DNA sequencing of the samples. Larger versions of panels **B** and **C** are provided in S1 and S2 Figs. Threshold cycle (C_t) and endpoint fluorescence values are given in S1 Table.

Mechanical breakage of the tissue sample emerged as a key variable for efficient and robust isolation of fDNA. Smaller fragments increase the surface area to volume ratio, allowing greater efficiency of proteinase K digestion. Also, the VFD removes solids from solution by centrifugation, and, therefore, breaking the tissue sample into small fragments improves the efficiency of VFD-mediated fDNA recovery. The formalin-fixed tissue was first frozen in liquid nitrogen then ground into small particles using a clean mortar and pestle. When extracted using traditional extraction approaches, the ground material yielded modest and inconsistent quantities of fDNA (Fig 2 and S1 Table). The lack of reproducibility suggested a

need for further mechanical breakdown of the tissue. Beadbeating and sonication increased the observed breakdown of tissue, but both methods failed to improve PCR product yields (data not shown). These failures may be due to the fact that sonication or the resultant heat generation through cavitation could introduce additional breaks in the DNA. However, we found that a second, mincing step with a sterile, new razorblade consistently improved fDNA recovery (Fig 4).

Also, the length of the amplified DNA proved critical for consistent recovery from formalin-fixed samples. In initial experiments, we amplified a 579 bp sequence of fDNA. The amplified DNA was the correct lobster sequence (Fig 2, Fig 4, and S2 Fig), but this long fragment proved difficult to amplify repeatedly. Thus, primers targeting a 183 bp sequence were used for all further experiments reported here. The shorter target amplicon decreased the chance of DNA fragmentation damaging the sequence and preventing PCR amplification. As expected, DNA sequencing quality in the non-VFD control reactions was difficult to monitor; only three out of thirteen cumulative control amplification reactions yielded fDNA detectable by gel electrophoresis.

The final parameter optimized for fDNA extraction was the rotational speed of the VFD. Previous studies with the VFD have demonstrated that acceleration of enzymatic catalysis occurs at rotational speeds between 5 and 9 krpm with the VFD at a 45° tilt angle [29]. Hypotheses attribute the enzymatic acceleration phenomena to two interconnected actions. First, the periodic change in the thickness of the thin fluid film present in the VFD, results in intense micro-mixing and high mass transfer. Second, the Faraday waves arising from this periodic change contribute to zones of high and low pressure within the reaction mixture and, accordingly, the enzyme present in solution. The pressure oscillations could increase substrate accessibility and removal of product from the enzyme active site, which also benefit from the high mass transfer of the VFD.

To identify optimal speeds for fDNA recovery, a systematic assessment of rotational speeds between 5 and 9 krpm at intervals of 1 krpm was conducted. Heterogeneity inherent to pulverized tissue imparts idiosyncratic and uncontrollable variables into this optimization and subsequent isolation of fDNA. Thus, the yields of fDNA, as determined by UV-Vis spectrophotometry, were sometimes inconsistent (It is well known that absorption of 260 nm is a proxy for DNA concentrations in solution, whereas the ratio of absorbances at 260:280 nm serves as an indicator of DNA purity). Though absorbance at 260 nm

increased for the VFD-processed samples at various rotational speeds relative to the non-VFD-processed negative control, the ratio of 260:280 nm absorbances was significantly <1.8 for all rotational speeds (Fig 3 and S3 Fig), which likely indicates inclusion of protein in the VFD-treated samples [35]. The DNA-associated, absorbance at 260 nm was greatest for the samples processed at 8 krpm. This rotational speed is within the VFD-based rate enhancement zone for reactions in aqueous solvents. That said, the highest, most consistently observed PCR yields were obtained for the rotational speed of the VFD of 7 krpm (which is discussed in more detail below; see Fig 2, Fig 4, S1 Fig, S2 Table, and S5 Fig).

Fig 3. Quantification for optimizing the VFD rotational speed for fDNA yields. After proteinase K and VFD treatment at the indicated speeds, **(A)** absorbance at 260 nm and **(B)** the ratio of absorbances 260:280 nm quantifies DNA and protein yields, respectively, with positive and negative controls. Full UV-vis spectra for these samples are shown in S3 Fig. SYBR Green I fluorescence-quantified **(C)** dsDNA concentration and **(D)** fold increase in dsDNA yield between non-VFD-processed and VFD-processed samples. The negative control indicates samples not subjected to VFD processing. Buffer only controls lacked lobster tissue. The positive controls included DNA that had not been formalin fixed. Additional controls demonstrated a conventional method and an intermediate method (processing time of the conventional method and temperature of the VFD-mediated method) used to process fixed and fresh tissue without the VFD. The error bars designate the standard deviation for sample measurements at the indicated condition (technical replicates, n = 3).

To more robustly quantify the concentration of double-stranded DNA (dsDNA) in the fDNA, a fluorescence-based DNA intercalating assay was performed using SYBR Green I dye. Samples were diluted to reach the dye's linear range, and the sample fluorescence (λ_{ex} 485 nm, λ_{em} 550 nm) was measured against a standard curve to determine the concentrations of the rescued dsDNA. The assay demonstrated that VFD-processed samples yielded 40 to 85% more dsDNA than the control, non-VFD processed sample (Fig 3). Thus, the VFD improved the rescue of dsDNA from formalin-fixed tissue compared to the non-VFD-treated negative control.

Moreover, and most importantly, DNA recovered via VFD-enhanced extraction were amenable to amplification via PCR and quantitative PCR (or qPCR). Post-VFD treatment, samples were purified and

318 concentrated with a Zymo™ DNA Clean & Concentrator Kit, which removes DNA polymerase inhibitors
319 and proteins [16, 36]. Using the fDNA samples and qPCR, we observed reproducible amplification of an
320 183 bp target amplicon from the gene encoding ATP synthase from samples processed with a VFD
321 rotating at either 6 or 7 krpm (Fig 4, Table 2). Paradoxically, these speeds had lower fDNA yields when
322 compared to the 8 krpm (as described above; see Fig 3a-b). However, the 8 krpm treated DNA would
323 amplify <50% of the time (n = 8, shown in S5 Fig). Though yields were greater, the more aggressive
324 treatment is apparently liberating other compounds that can be problematic to PCR amplification (this
325 phenomena has also been seen in algal cells and more) [37]. Additionally, it is plausible that greater VFD
326 rotational speeds could result in both greater yields and further fragmentation of fDNA, which could
327 explain the higher DNA concentrations, but lower amplification efficacy [16].

328 From the formalin-fixed samples, the 7 krpm VFD-processed DNA sample yielded the highest
329 levels of fDNA amplification as measured by qPCR. For example, on average 94±4% of the endpoint
330 positive control fluorescence signal was obtained with a low threshold cycle (C_t) value, averaging
331 35.4±0.7 cycles (S2 Table). Comparatively, the no template controls (NTC) did not surpass the threshold
332 in two of three trials, and their fluorescence averaged 10±10% of the endpoint positive control
333 fluorescence signal (S2 Table). The lower C_t values demonstrated a greater yield of DNA. Furthermore,
334 the PCR product of the fDNA rescued from the 7 krpm VFD-processing condition could readily be
335 sequenced via Sanger sequencing. This sequence exhibited 99.5% homology to the expected sequence
336 [38]. The 6 krpm-processed sample offered a decreased yield of amplified DNA. Notably, PCR
337 amplification failed for the non-VFD-treated sample.

338 The data herein illustrate the efficacy of our new VFD-enabled method for fDNA recovery. We
339 have demonstrated the successful amplification of fDNA from biological specimens treated with
340 formaldehyde. We have also shown that fDNA can be used to great effect with appropriately designed
341 qPCR assays. Therefore, we are optimistic that this method presents a potentially valuable method for
342 increasing the throughput of fDNA recovery. Increasing the rate at which fDNA can be recovered is a
343 timely pursuit, as there are tens of millions of formalin-fixed samples stored in museums around the world.
344 These organisms provide a “time capsule” of sorts, revealing the genomic adaptations of organisms to a
345 pre-industrial world. Indeed, museum specimens may become our first tool for understanding the extent

346 to which anthropogenic factors are shaping our biosphere. It is also important to note that fDNA is,
347 unfortunately, subject to irreparable damage and fragmentation, and any PCR-based amplification from
348 such samples will always be highly dependent on rescue conditions, primer design (especially amplicon
349 length), tissue mincing, and other factors. The VFD method here may help investigators tap into this
350 enormous genetic repository.
351

352 **Fig 4. Amplification of an 183-bp fDNA target from the ATP synthase gene of the lobster**
353 **mitochondrial genome. (A)** Quantitative PCR and **(B)** agarose DNA gel electrophoresis identified 7
354 krpm as the optimal VFD rotational speed for qPCR amplification. Threshold cycle and endpoint
355 fluorescence values are provided in S2 Table. The variable-rotational speed PCR reactions were
356 compared to a no template control (NTC), a fresh lobster DNA positive control (+), and a non-VFD-
357 processed negative control (–). **(C)** The 7 krpm VFD-processed qPCR product (*) was subjected to
358 Sanger sequencing; a mutation (G2728A, GenBank No. HQ402925) was observed (highlighted).

359 # Availability

360 Genbank is a collection of gene and genome sequences in the National Center for Biotechnology
361 Information database (https://www.ncbi.nlm.nih.gov/genbank/).

362 # Author information

363 ## ORCID

364 CA Totoiu: 0000-0001-9839-8894
365 JM Phillips: 000-0001-5673-0879
366 AT Reese: 0000-0001-9004-9470
367 S Majumdar: 0000-0001-6738-2267
368 PR Girguis: 0000-0002-3599-8160
369 CL Raston: 0000-0003-4753-0079
370 GA Weiss: 0000-0003-0296-9846

Acknowledgements

We gratefully thank UCI's Department of Molecular Biology and Biochemistry for access to the qPCR thermocycler.

Funding

This work was supported by the National Human Genome Research Institute (NHGRI) of the National Institutes of Health [1R01HG009188-01 to G.A.W.]; the Star Family Foundation Challenge for Promising Scientific Research [to A.T.R. and P.R.G.]; and the University of California at Irvine's Undergraduate Research Opportunities Program and Summer Undergraduate Research Program [to C.A.T.]. Funding for open access charge: NHGRI.

Conflict of interest

The authors declare no competing financial interests. Debut Biotechnology, a company co-founded by Dr. Weiss, has licensed the VFD technology for different applications.

References

[1] Holmes,M.W., Hammond,T.T., Wogan,G.O., Walsh,R.E., LaBarbera,K., Wommack,E.A., Martins,F.M., Crawford,J.C., Mack,K.L., Bloch,L.M. and Nachman,M.W. (2016) Natural history collections as windows on evolutionary processes. *Mol. Ecol.*, **25**, 864–881.

[2] Suarez,A.V. and Tsutsui,N.D. (2004) The value of museum collections for research and society. *BioScience*, **54**, 66–74.

[3] Pergams,O.R.W. and Nyberg,D. (2001) Museum collections of mammals corroborate the exceptional decline of prairie habitat in the Chicago region. *J. Mammal.*, **82(4)**, 984-992.

[4] Lens,L., Van Dongen,S., Norris,K., Githiru,M. and Matthysen,E. (2002) Avian persistence in fragmented rainforest. *Science*, **298(5596)**, 1236-1238.

[5] Carroll,S.P. and Boyd,C. (1992) Host race radiation in the soapberry bug: natural history with the history. *Evolution*, **46(4)**, 1052-1069.

[6] Greer,C.E., Lund,J.K. and Manos,M.M. (1991) PCR amplification from paraffin-embedded tissues: recommendations on fixatives for long-term storage and prospective studies. *PCR Methods Appl.*, **1(1)**, 46-50.

[7] Hykin,S.M., Bi,K. and Mcguire,J.A. (2015) Fixing formalin: a method to recover genomic-scale DNA sequence data from formalin-fixed museum specimens using high-throughput sequencing. *PLoS ONE*, **10(10)**, e0141579.

[8] Mikheyev,A.S., Tin,M.M.Y., Arora,J. and Seeley,T.D. (2015) Museum samples reveal rapid evolution by wild honey bees exposed to a novel parasite. *Nat. Commun.*, **6**, 7991.

[9] Spurgin,L.G., Wright,D.J., van der Velde,M., Collar,N.J., Komdeur,J., Burke,T. and Richardson,D.S. (2014) Museum DNA reveals the demographic history of the endangered Seychelles warbler. *Evol. Appl.*, **7(9)**, 1134-1143.

[10] Fox,C.H., Johnson,F.B., Whiting,J. and Roller,P.P. (1985) Formaldehyde fixation. *J. Histochem. Cytochem.*, **33(8)**, 845-853.

[11] Do,H. and Dobrovic,A. (2015) Sequence artifacts in DNA from formalin-fixed tissues: causes and strategies for minimization. *Clin. Chem.*, **61(1)**, 64-71.

[12] Suzuki,T., Ohsumi,S. and Makino,K. (1994) Mechanistic studies on depurination and apurinic site chain breakage in oligodeoxyribonucleotides. *Nucleic Acids Res.*, **22(23)**, 4997-5003.

[13] Lindahl,T. and Andersson,A. (1972) Rate of chain breakage at apurinic sites in double-stranded deoxyribonucleic acid. *Biochemistry*, **11(19)**, 3618-3623.

[14] Ludyga,N., Grunwald,B., Azimzadeh,O., Englert,S., Hofler,H., Tapio,S. and Aubele,M. (2012) Nucleic acids from long-term preserved FFPE tissues are suitable for downstream analyses. *Virchows Arch.*, **460**, 131-140.

[15] Didelot,A., Kotsopoulos,S.K., Lupo,A., Pekin,D., Li,X., Atochin,I., Srinivasan,P., Zhong,Q., Olson,J., Link,D.R., Laurent-Piug,P., Blons,H., Hutchison,J.B. and Taly,V. (2013) Multiplex picoliter-droplet digital PCR for quantitative assessment of DNA integrity in clinical samples. *Clin. Chem.*, **59(5)**, 815-823.

[16] Dietrich,D., Uhl,B., Sailer,V., Holmes,E.E., Jung,M., Meller,S. and Kristiansen,G. (2013) Improved PCR performance using template DNA from formalin-fixed and paraffin-embedded tissues by overcoming PCR inhibition. *PloS ONE*, **8(10)**, e77771.

[17] Dutta,S., Chowdhury,G. and Gates,K.S. (2007) Interstrand cross-links generated by abasic sites in duplex DNA. *J. Am. Chem. Soc.*, **129(7)**, 1852-1853.

[18] Vesnaver,G., Chang,C.N., Eisenberg,M., Grollman,A.P. and Breslauer,K.J. (1989) Influence of abasic and anucelosidic sites on the astability, conformation, and melting behavior of a DNA duplex: correlations of thermodynamic and structural data. *Proc. Natl. Acad. Sci. U.S.A.*, **86(10)**, 3614-3618.

[19] Kennedy-Darling,J. and Smith,L.M. (2014) Measuring the formaldehyde protein−DNA cross-link reversal rate. *Anal. Chem.*, **86(12)**, 5678−5681.

[20] Hoffman,E.A., Frey,B.L., Smith,L.M. and Auble,D.T. (2015) Formaldehyde crosslinking: A tool for the study of chromatin complexes. *J. Biol. Chem.*, **290(44)**, 26404-26411.

[21] Kavli,B., Otterlei,M., Slupphaug,G. and Krokan,H.E. (2007) Uracil in DNA-general mutagen, but normal intermediate in acquired immunity. *DNA Repair (Amst.)*, **6(4)**, 505-516.

[22] Yazawa,K., Sugahara,M., Yutani,K., Takehira,M. and Numata,K. (2016) Derivatization of proteinase K with heavy atoms enhances its thermal stability. *ACS Catal.*, **6(5)**, 3036−3046.

[23] Goelz,S.E., Hamilton,S.R. and Vogelstein,B. (1985) Purification of DNA from formaldehyde fixed and paraffin embedded human tissue. *Biochem. Biophys. Res. Commun.*, **130(1)**, 118-126.

[24] Butler,J.M., Johnson,J.E. and Boone,W.R. (2013) The heat is on: room temperature affects laboratory equipment–an observational study. *J. Assist. Reprod. Genet.*, **30(10)**, 1389-1393.

[25] Campos,P.F. and Gilbert,T.M.P. (2012) DNA extraction from formalin-fixed material. In Shapiro,B. and Hofreiter,M. (eds) *Ancient DNA. Methods in Molecular Biology (Methods and Protocols)*. Humana Press, Vol. 840.

[26] Munchel,S., Hoang,Y., Zhao,Y., Cottrell,J., Klotzle,B., Godwin,A.K., Koestler,D., Beyerlein,P., Fan,J., Bibikova,M. and Chien,J. (2015) Targeted or whole genome sequencing of formalin fixed tissue samples: potential applications in cancer genomics. *Oncotarget.*, **6(28)**, 25943-25961.

[27] Hassani,A. and Khan,G. (2015) A simple procedure for the extraction of DNA from long-term formalin-preserved brain tissues for the detection of EBV by PCR. *Exp. Mol. Pathol.*, **99(3)**, 558-563.

[28] Paireder,S., Werner,B., Bailer,J., Werther,W., Schmid,E., Patzek,B. and Cichna-Markl,M. (2013) Comparison of protocols for DNA extraction from long-term preserved formalin fixed tissues. *Anal. Biochem.*, **439(2)**, 152-160.

[29] Britton,J., Meneghini,L.M., Raston,C.L. and Weiss,G.A. (2016) Accelerating enzymatic catalysis using vortex fluidics. *Angew. Chem. Int. Ed.*, **55(38)**, 11387-11391.

[30] Yuan,T.Z., Ormonde,C.F.G., Kudlacek,S.T., Kunche,S., Smith,J.N., Brown,W.A., Pugliese,K.M., Olsen,T.J., Iftikhar,M., Raston,C.L. and Weiss,G.A. (2015) Shear stress-mediated refolding of proteins from aggregates and inclusion bodies. *ChemBioChem.*, **16**, 393-396.

[31] Hu,M., Jex,A.R., Campbell,B.E. and Gasser,R.B. (2007) Long PCR amplification of the entire mitochondrial genome from individual helminths for direct sequencing. *Nat. Protoc.*, **2(10)**, 2339-2344.

[32] Britton,J., Stubbs,K.A., Weiss,G.A. and Raston,C.L. (2017) Vortex fluidic chemical transformations, *Chem. Eur. J.*, **23**, 13270-13278.

[33] Luo,X., Al-Antaki,A.H.M., Harvey,D.P., Ruan,Y., He,S., Zhang,W. and Raston,C.L. (2018) Vortex fluidic mediated synthesis of macroporous bovine serum albumin-based microspheres. *ACS Appl. Mater. Interfaces.*, **10**, 27224-27232.

[34] Leggate,J., Allain,R., Isaac,L. and Blais,B.W. (2006) Microplate fluorescence assay for the quantification of double stranded DNA using SYBR Green I dye. *Biotechnol. Lett.*, **28(19)**, 1587-1594.

[35] Glasel,J.A. (1995) Validity of nucleic acid purities monitored by 260 nm/280nm absorbance ratios. *BioTechniques.*, **18(1)**, 62-63.

469 [36] Hu,Q., Liu,Y., Yi,S. and Huang,D. (2015) A comparison of four methods for PCR inhibitor
470 removal. *Forensic Sci. Int. Genet*., **16**, 94-97.

471 [37] Sitepu,E.K., Corbin,K., Luo,X., Pye,S.J., Tang,Y., Leterme,S.C., Heimann,K., Raston,C.L. and
472 Zhang,W. (2018), *Bioresour Technol*., **266**, 488-497.

473 [38] Kim,S., Lee,S.H., Park,M.H., Choi,H.G., Park,J.K. and Min,G.S. (2011) The complete
474 mitochondrial genome of the American lobster, Homarus americanus (Crustacea, Decapoda)
475 *Mitochondrial DNA*, **22(3)**, 47-49.

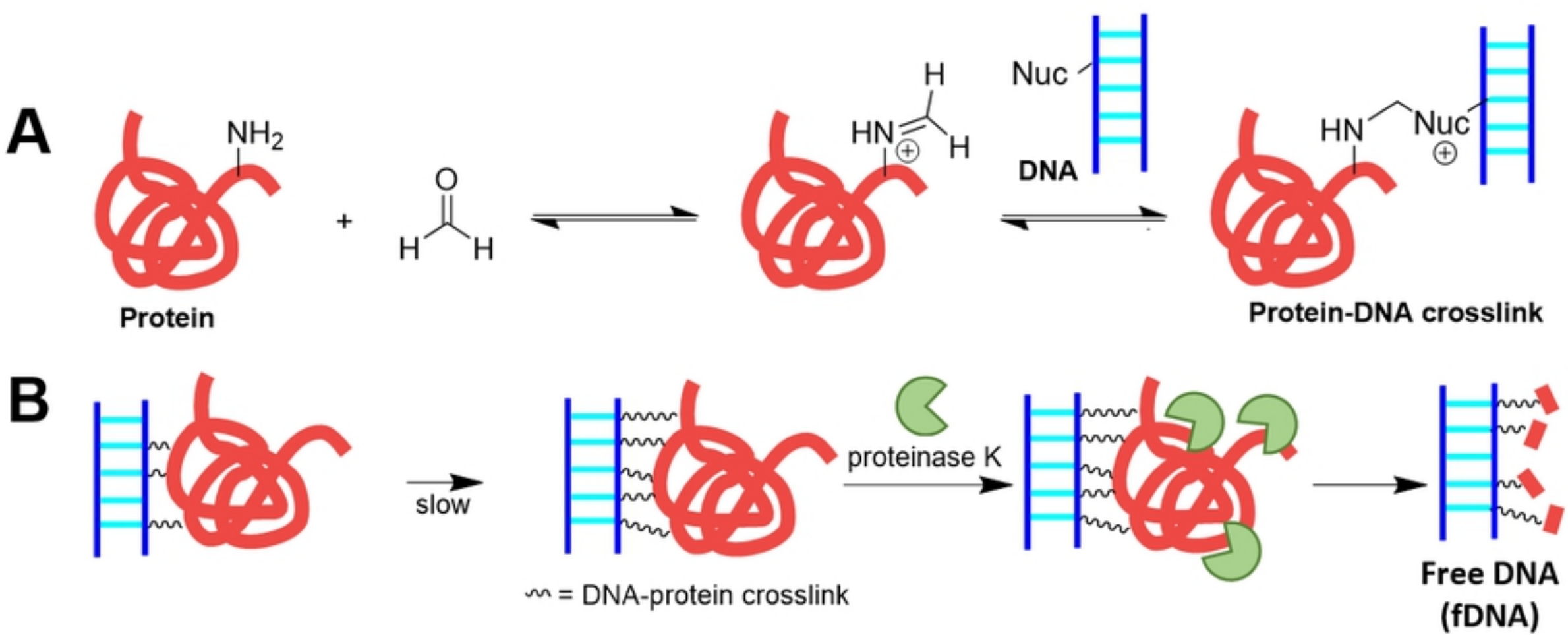

A
NH2
Protein
+
O
H
H
H
HN
H
Nuc
DNA
HN
Nuc
Protein-DNA crosslink
B
slow
proteinase K
= DNA-protein crosslink
Free DNA
(fDNA)

Figure 1

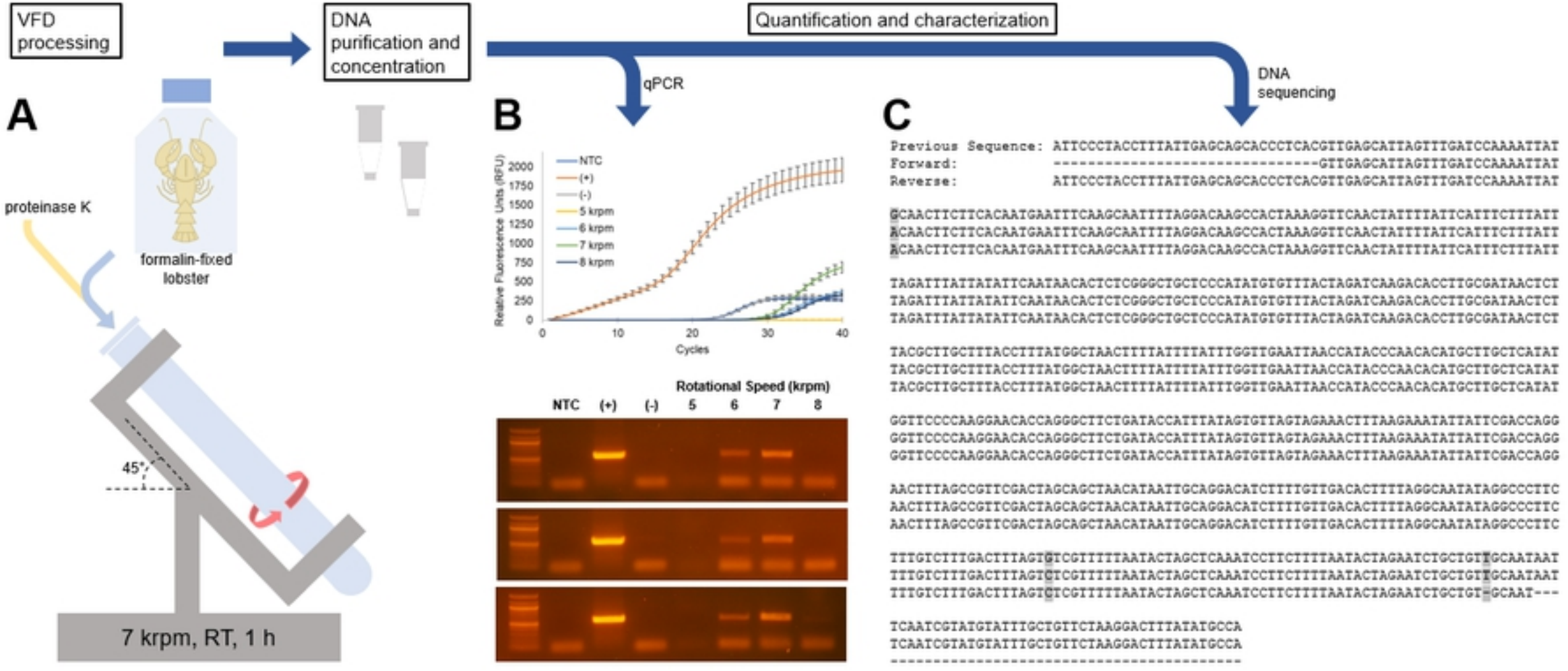

VFD processing
DNA purification and concentration
Quantification and characterization
qPCR
DNA sequencing
A
proteinase K
formalin-fixed lobster
45°
7 krpm, RT, 1 h
B
Relative Fluorescence Units (RFU)
Cycles
NTC
(+)
(-)
5 krpm
6 krpm
7 krpm
8 krpm
Rotational Speed (krpm)
NTC (+) (-) 5 6 7 8
C
Previous Sequence: ATTCCCTACCTTTATTGAGCAGCACCCTCACGTTGAGCATTAGTTTGATCCAAAATTAT
Forward: ----------GTTGAGCATTAGTTTGATCCAAAATTAT
Reverse: ATTCCCTACCTTTATTGAGCAGCACCCTCACGTTGAGCATTAGTTTGATCCAAAATTAT
GCAACTTCTTCACAATGAATTTCAAGCAATTTTAGGACAAGCCACTAAAGGTTCAACTATTTTATTCATTTCTTTATT
ACAACTTCTTCACAATGAATTTCAAGCAATTTTAGGACAAGCCACTAAAGGTTCAACTATTTTATTCATTTCTTTATT
ACAACTTCTTCACAATGAATTTCAAGCAATTTTAGGACAAGCCACTAAAGGTTCAACTATTTTATTCATTTCTTTATT
TAGATTTATTATATTCAATAACACTCTCGGGCTGCTCCCATATGTGTTTACTAGATCAAGACACCTTGCGATAACTCT
TAGATTTATTATATTCAATAACACTCTCGGGCTGCTCCCATATGTGTTTACTAGATCAAGACACCTTGCGATAACTCT
TAGATTTATTATATTCAATAACACTCTCGGGCTGCTCCCATATGTGTTTACTAGATCAAGACACCTTGCGATAACTCT
TACGCTTGCTTTACCTTTATGGCTAACTTTTATTTTATTTGGTTGAATTAACCATACCCAACACATGCTTGCTCATAT
TACGCTTGCTTTACCTTTATGGCTAACTTTTATTTTATTTGGTTGAATTAACCATACCCAACACATGCTTGCTCATAT
TACGCTTGCTTTACCTTTATGGCTAACTTTTATTTTATTTGGTTGAATTAACCATACCCAACACATGCTTGCTCATAT
GGTTCCCCAAGGAACACCAGGGCTTCTGATACCATTTATAGTGTTAGTAGAAACTTTAAGAAATATTATTCGACCAGG
GGTTCCCCAAGGAACACCAGGGCTTCTGATACCATTTATAGTGTTAGTAGAAACTTTAAGAAATATTATTCGACCAGG
GGTTCCCCAAGGAACACCAGGGCTTCTGATACCATTTATAGTGTTAGTAGAAACTTTAAGAAATATTATTCGACCAGG
AACTTTAGCCGTTCGACTAGCAGCTAACATAATTGCAGGACATCTTTTGTTGACACTTTTAGGCAATATAGGCCCTTC
AACTTTAGCCGTTCGACTAGCAGCTAACATAATTGCAGGACATCTTTTGTTGACACTTTTAGGCAATATAGGCCCTTC
AACTTTAGCCGTTCGACTAGCAGCTAACATAATTGCAGGACATCTTTTGTTGACACTTTTAGGCAATATAGGCCCTTC
TTTGTCTTTGACTTTAGTGTCGTTTTTAATACTAGCTCAAATCCTTCTTTTAATACTAGAATCTGCTGTTGCAATAAT
TTTGTCTTTGACTTTAGTGTCGTTTTTAATACTAGCTCAAATCCTTCTTTTAATACTAGAATCTGCTGTTGCAATAAT
TTTGTCTTTGACTTTAGTGTCGTTTTTAATACTAGCTCAAATCCTTCTTTTAATACTAGAATCTGCTGT-GCAAT---
TCAATCGTATGTATTTGCTGTTCTAAGGACTTTATATGCCA
TCAATCGTATGTATTTGCTGTTCTAAGGACTTTATATGCCA

Figure 2

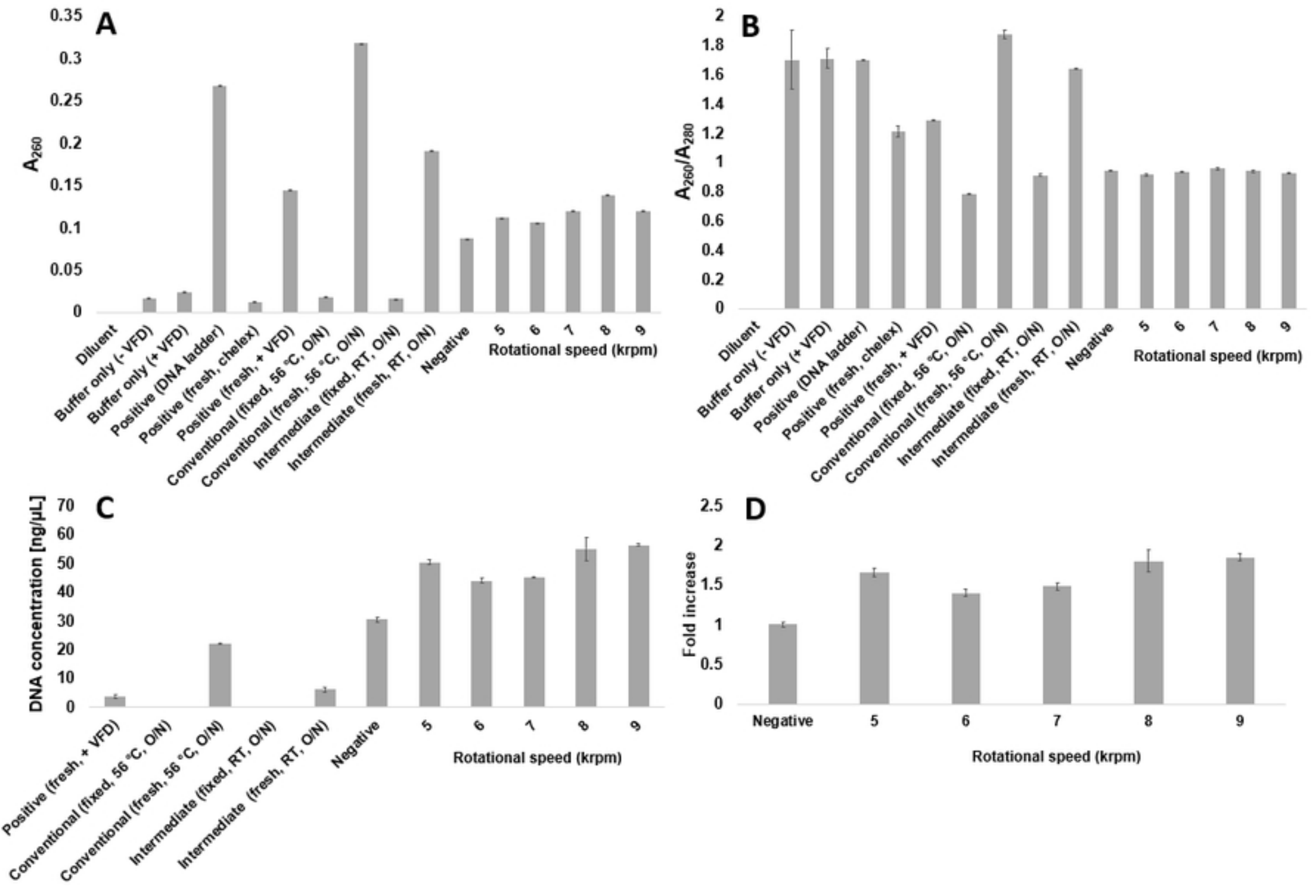
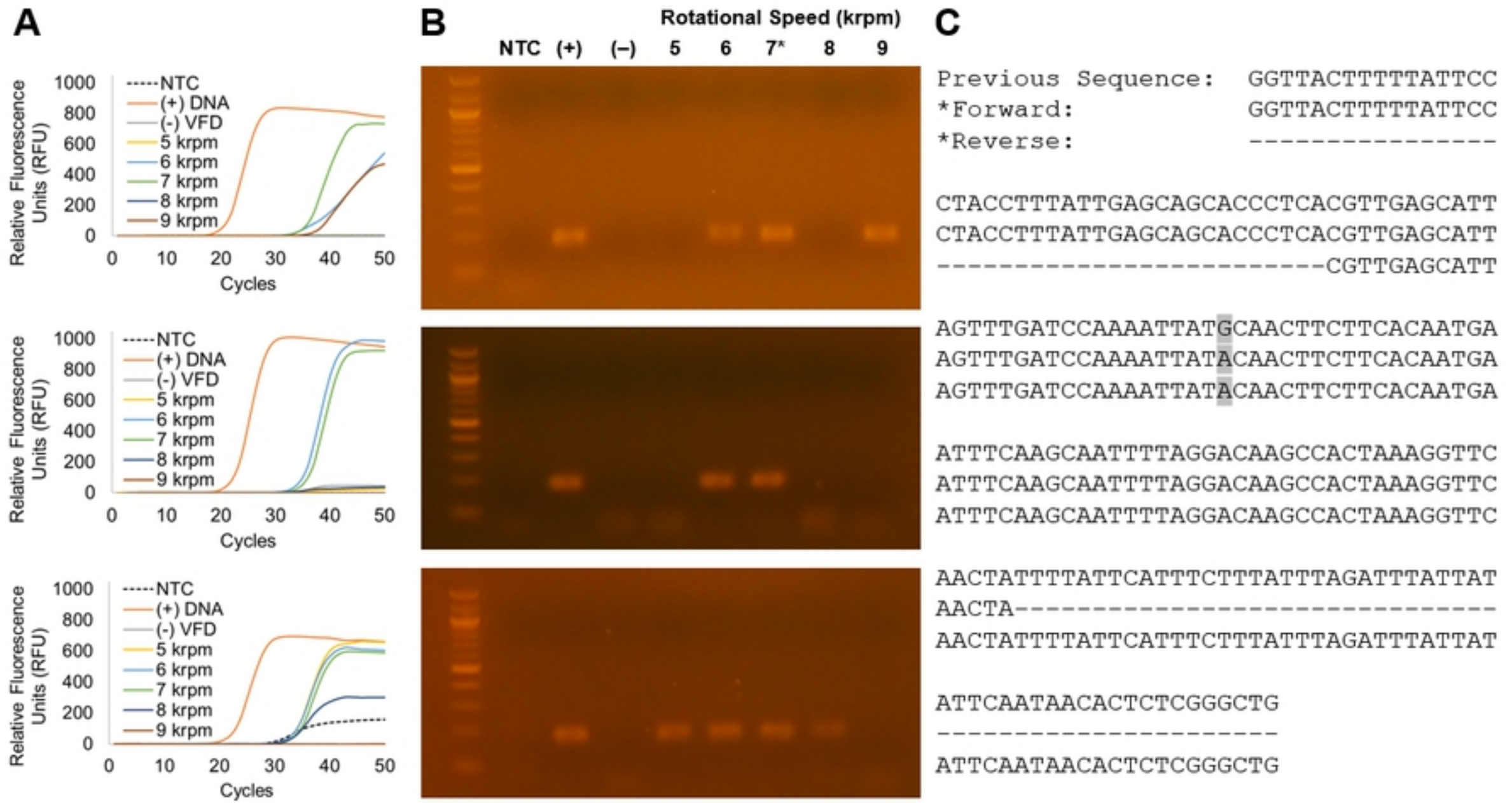

A
A_{260}
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Diluent
Buffer only (- VFD)
Buffer only (+ VFD)
Positive (DNA ladder)
Positive (fresh, chelex)
Positive (fresh, + VFD)
Conventional (fixed, 56 °C, O/N)
Conventional (fresh, 56 °C, O/N)
Intermediate (fixed, RT, O/N)
Intermediate (fresh, RT, O/N)
Negative
5
6
7
8
9
Rotational speed (krpm)
B
A_{260}/A_{280}
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Diluent
Buffer only (- VFD)
Buffer only (+ VFD)
Positive (DNA ladder)
Positive (fresh, chelex)
Positive (fresh, + VFD)
Conventional (fixed, 56 °C, O/N)
Conventional (fresh, 56 °C, O/N)
Intermediate (fixed, RT, O/N)
Intermediate (fresh, RT, O/N)
Negative
5
6
7
8
9
Rotational speed (krpm)
C
DNA concentration [ng/µL]
70
60
50
40
30
20
10
0
Positive (fresh, + VFD)
Conventional (fixed, 56 °C, O/N)
Conventional (fresh, 56 °C, O/N)
Intermediate (fixed, RT, O/N)
Intermediate (fresh, RT, O/N)
Negative
5
6
7
8
9
Rotational speed (krpm)
D
Fold increase
2.5
2
1.5
1
0.5
0
Negative
5
6
7
8
9
Rotational speed (krpm)

Figure 3


A
B
C
Rotational Speed (krpm)
NTC (+) (–) 5 6 7* 8 9
Relative Fluorescence Units (RFU)
Cycles
NTC
(+) DNA
(-) VFD
5 krpm
6 krpm
7 krpm
8 krpm
9 krpm
Previous Sequence: GGTTACTTTTTATTCC
*Forward: GGTTACTTTTTATTCC
*Reverse: ----------
CTACCTTTATTGAGCAGCACCCTCACGTTGAGCATT
CTACCTTTATTGAGCAGCACCCTCACGTTGAGCATT
----------CGTTGAGCATT
AGTTTGATCCAAAATTATGCAACTTCTTCACAATGA
AGTTTGATCCAAAATTATACAACTTCTTCACAATGA
AGTTTGATCCAAAATTATACAACTTCTTCACAATGA
ATTTCAAGCAATTTTAGGACAAGCCACTAAAGGTTC
ATTTCAAGCAATTTTAGGACAAGCCACTAAAGGTTC
ATTTCAAGCAATTTTAGGACAAGCCACTAAAGGTTC
AACTATTTTATTCATTTCTTTATTTAGATTTATTAT
AACTA----------
AACTATTTTATTCATTTCTTTATTTAGATTTATTAT
ATTCAATAACACTCTCGGGCTG

ATTCAATAACACTCTCGGGCTG

Figure 4