Published in final edited form as:

Curr Protoc Microbiol. ; 33: 1E.8.1–1E.8.16. doi:10.1002/9780471729259.mc01e08s37.

Isolating Viral and Host RNA Sequences from Archival Material and Production of cDNA Libraries for High-Throughput DNA Sequencing

Yongli Xiao¹, Zong-Mei Sheng¹, and Jeffery K. Taubenberger¹

¹Viral Pathogenesis and Evolution Section, Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland

Abstract

The vast majority of surgical biopsy and post-mortem tissue samples are formalin-fixed and paraffin-embedded (FFPE), but this process leads to RNA degradation that limits gene expression analysis. As an example, the viral RNA genome of the 1918 pandemic influenza A virus was previously determined in a 9-year effort by overlapping RT-PCR from post-mortem samples. Using the protocols described here, the full genome of the 1918 virus at high coverage was determined in one high-throughput sequencing run of a cDNA library derived from total RNA of a 1918 FFPE sample after duplex-specific nuclease treatments. This basic methodological approach should assist in the analysis of FFPE tissue samples isolated over the past century from a variety of infectious diseases.

Keywords

RNA; cDNA; high-throughput sequencing; library; influenza; polymerase chain reaction

INTRODUCTION

Molecular diagnostics approaches for a wide variety of infectious diseases have been successfully applied to archival fixed tissue biopsy and autopsy samples for over two decades (O'Leary 2003). Polymerase chain reaction (PCR)-based or Reverse Transcription PCR (RT-PCR)-based approaches have been readily applied to microbial diagnosis because of their ease, sensitivity, and specificity. These types of assays are very useful for clinical diagnosis of a suspected pathogen, but of course require at least partial pathogen genome sequence target information to design appropriate primers, probes, and assay conditions.

In the last decade, high-throughput DNA sequencing platforms, or next generation sequencing, which can determine millions of bases (megabases) of DNA sequence per run (Service 2006), have provided researchers, diagnostic laboratories, and public health officials with a new and powerful tool. High-throughput DNA sequencing strategies can be designed using target sequence-dependent strategies, but significantly, sequence target-independent high-throughput sequencing methodologies allow for the identification of novel viral genomes using metagenomic approaches (Bibby 2013; Firth and Lipkin 2013;

Baldwin, Feldman et al. 2014). These approaches have the potential to transform pathogenesis research, surveillance, and clinical microbiological diagnostics. For identification or characterization of RNA viral genomes, most of these strategies depend, however, on high-quality viral RNA.

Unfortunately, most biopsy or autopsy tissue samples are formalin-fixed and paraffin-embedded (FFPE). An enormous number of FFPE archival human and veterinary tissue samples from the end of the 19th Century (Fox, Johnson et al. 1985) to the present are stored in hospitals, tissue banks, and laboratories. These samples have the potential to provide a wealth of novel information in retrospective molecular genetic and/or metagenomic studies of diseased tissues (Tang, David et al. 2009). Different analysis platforms, including microarray (April, Klotzle et al. 2009; Mittempergher, de Ronde et al. 2011) and high-throughput sequencing (Schweiger, Kerick et al. 2009; Beck, Weng et al. 2010; Wood, Belvedere et al. 2010) have been applied to FFPE samples. A significant challenge to the metagenomic analysis of these archival samples, however, is the degradation and modification of RNA that occurs during fixation and embedding (Krafft, Duncan et al. 1997; Masuda, Ohnishi et al. 1999; Evers, Fowler et al. 2011; Evers, He et al. 2011).

Recently, we applied a high-throughput DNA sequencing metagenomic approach to FFPE human autopsy lung tissue samples from the 1918 ‘Spanish’ influenza pandemic (Xiao, Kash et al. 2013). Different research or diagnostic questions will require development and implementation of different strategies, but this study serves as a representative methodology to perform a metagenomic analysis from fragmented RNA isolated from archival material nearly 100 years old. The protocol presented here describes the construction of deep sequencing libraries suitable for the Illumina platform from RNA isolated from archival FFPE tissues for the enrichment and detection of viral RNA and host mRNA. It contains three parts: Basic Protocol 1 describes the steps for the effective isolation of total RNA from archival FFPE samples; Basic Protocol 2 describes the steps of preparing a total RNA sequencing library from isolated RNA for Illumina high throughput sequencing; Basic Protocol 3 describes the normalization steps for the prepared sequencing library using duplex-specific thermostable nuclease (DSN) to decrease the amount of cDNA derived from rRNA in order to enrich for cDNAs derived from viral RNA and host mRNA in the library.

BASIC PROTOCOL 1: ISOLATION OF RNA FROM ARCHIVAL FIXED TISSUE SAMPLES

RNA can be routinely isolated from the vast majority of formalin-fixed and paraffin-embedded (FFPE) tissue specimens. To isolate RNA, tissue sections are first deparaffinized with Hemo-De. The tissue is then washed in ethanol to remove residual Hemo-De or CitriSolv. Cells are lysed in a detergent buffer, which also inhibits endogenous RNases. RNA isolation involves extraction with TRIzol LS followed by isopropanol precipitation. This procedure results in purification of the total RNA away from the bulk of the DNA and yields a preparation suitable for RT-PCR assays or cDNA library production.

Materials

CitriSolv
Absolute ethanol
Proteinase K (20 mg/ml in molecular-grade water), stored at -20°C
Extraction buffer (see recipe)
3M Sodium acetate
Glycogen, stored at -20°C
Tris-EDTA buffer, pH 8 (TE)
Dulbecco's Phosphate-buffered saline (PBS)
75% Ethanol
Isopropanol
TRIzol LS, stored at 4°C
Phenol/Chloroform/Isoamyl alcohol, stored at 4°C
Chloroform
DEPC treated water
Microcentrifuge tubes (1.5 ml)
Aerosol resistant pipette tips (ART tips)
Disposable gloves
Sterile conical Polypropylene tubes (15 and 50 ml)
Pipettors
Vortexer
Microcentrifuge
Oven (55°C)
Water bath (55°C)
Spectrophotometer
Table-top centrifuge

Deparaffinization of tissue sections and initial extraction

- 1 Add 800 μl of CitriSolv to an RNase-free microcentrifuge tube containing three-to-six 6 μm sections of tissue. Vortex at full speed for 5 sec. Add 400 μl ethanol. Vortex at full speed for 5 sec. Spin in a microcentrifuge at full speed ($\sim 14,000$ rpm) for 5 min. If residual paraffin is visible, repeat this step.
- 2 Decant the liquid carefully and add 800 μl ethanol. Vortex for 5 sec. Centrifuge at full speed for 5 min. Decant the liquid.

- 3 Dry pellet in a 55°C oven for approximately 5 min or until air dry for at least 15 min at room temperature. Do not over dry.
- 4 Determine total amount of extraction buffer needed (250µl/sample). Remove proteinase K stock from freezer and thaw at room temperature. Tap proteinase K to mix. Vortex extraction buffer to mix. Add 45µl proteinase K to 1.5 ml extraction buffer and mix by tapping.
- 5 Add extraction buffer (250µl) to specimen tubes and vortex a few seconds at slow speed. Place in 55°C waterbath for 4 hr. to overnight.

Isolation of RNA

- 6 Add 750µl TRIzol LS to an RNase-free microcentrifuge tube containing 250µl of tissue in extraction buffer. Vortex at medium speed for 5 sec to mix thoroughly. Incubate at RT for 5–10 min. Add 200µl chloroform and shake by hand for 15–20 sec. Incubate at RT 5–10 min. Centrifuge at 12,000 rpm in a microcentrifuge for 10 min.
- 7 Transfer upper aqueous layer to a fresh 1.5 ml microcentrifuge tube containing 1.5µl glycogen (30µg). Add 500µl isopropanol and mix by inversion. Incubate on ice for at least 10 min.
- 8 Collect precipitate by centrifugation at 12,000 rpm for 10 min.
- 9 Wash pellet with 1.0 ml 75% ethanol, centrifuge at 9,000 rpm for 5 min. Decant supernatant.
- 10 Collect residual liquid at the bottom of the tube by centrifuging for a few seconds. Remove liquid with a pipet. Allow any remaining liquid to evaporate by leaving tube inverted on a Kimwipe for 10–15 min. Add 35µl of DEPC treated water. Incubate in 55°C waterbath for 10 min to allow RNA to resuspend. Mix gently and store at –70°C.

DNase treatment of RNA before making cDNA

- 11 DNase treat RNA by adding 1µl 10x DNase reaction Buffer, 1µl DNase I (Amplification grade, 1U/µl) and DEPC treated water to RNA, to a total reaction of 10µl.
- 12 Incubate at room temperature for 15 min, then add 1µl 25mM EDTA.
- 13 Stop reaction by incubation at 65 °C for 10 min.

BASIC PROTOCOL 2: SEQUENCING LIBRARY PREPARATION

Total RNA-seq libraries can be produced from RNA isolated from FFPE tissues for subsequent analysis by high-throughput sequencing. This protocol includes several main points: 1) Checking the input RNA quality and quantity using the Agilent RNA 6000 Nano kit on a Bioanalyzer. Typically, RNA isolated from archival FFPE samples is highly degraded with RNA peaks of around 100bp, which eliminates the need to shear input RNA during library construction. 2) Making cDNA, repairing ends, adenylating 3' ends, ligating

the adapters, enriching the cDNA templates, and validating the final library using Agilent High Sensitivity DNA kit on Bioanalyzer.

Materials

Agilent 2100 Bioanalyzer (includes IKA vortexer)
Agilent RNA 6000 Nano Kit
Agilent High Sensitivity DNA Kit
TruSeq® RNA Sample Preparation kit v2
SuperScript II
AMPure XP beads
Nuclease-free water
Heat block
Magnetic stand
Thermal cycler
Phusion Polymerase
25 mM dNTPs

Check input RNA quality and quantity using Agilent RNA 6000 Nano Kit

Preparing the RNA Ladder

1. Briefly microcentrifuge (“spin down”) the tube containing the RNA ladder and pipette in an RNase-free vial.
2. Heat denature the ladder for 2 min at 70 °C.
3. Immediately cool the vial on ice.
4. Prepare aliquots in 0.5 mL RNase-free vials with the required amount for typical daily use.
5. Store aliquots at –70 °C. After initial heat denaturation, the frozen aliquots should not require repeated heat denaturation.

Typically, the RNA ladder can be stored at –70 °C up to one year.

6. Before use, thaw ladder aliquots on ice (avoid extensive warming).

Preparing the Gel

1. Pipette 550 µl of RNA gel matrix (red) into a spin filter.
2. Centrifuge at 1500 g ± 20 % for 10 min at room temperature.
3. Aliquot 65 µl filtered gel into 0.5 mL RNase-free microcentrifuge tubes. Use filtered gel within 4 weeks. Store at 4 °C.

Preparing the Gel-Dye Mix

1. Allow the RNA dye concentrate (blue) to equilibrate to room temperature for 30 min.
2. Vortex RNA dye concentrate (blue) for 10 s, spin down and add 1 μ L of dye into a 65 μ L aliquot of filtered gel.
3. Vortex solution well. Spin tube at 13000 g for 10 min at room temperature. Use prepared gel- dye mix within one day.

Loading the Gel-Dye Mix

1. Put a new RNA chip on the chip priming station.
2. Pipette 9 μ L of gel- dye mix in the well marked with the symbol 'G.'
3. Make sure that the plunger is positioned at 1 mL and then close the chip priming station.
4. Press plunger until it is held by the clip.
5. Wait for exactly 30 sec. then release the clip.
6. Wait for 5 sec. and slowly pull back plunger to the 1 mL position.
7. Open the chip priming station and pipette 9 μ L of gel- dye mix in the wells with the symbol 'G.'
8. Discard the remaining gel-dye mix.

Loading the Marker:

Pipette 5 μ L of RNA marker (with green cap) in all 12 sample wells and in the well marked with the ladder symbol, which is used as nucleotide migration reference for different wells and the ladder well.

Loading the Ladder and Samples

1. Pipette 1 μ L of prepared ladder in well marked with the ladder symbol.
2. Pipette 1 μ L of sample in each of the 12 sample wells.
3. Pipette 1 μ L of RNA Marker (green) in each unused sample well.
4. Put the chip horizontally in the IKA vortexer and vortex for 1 min at 2400 rpm.
5. Run the chip in the Agilent 2100 Bioanalyzer instrument within 5 min.

The amount of the total FFPE RNA input should be 100–300ng. The typical FFPE total RNA profile on Bioanalyzer is shown in Figure 1.

Construction of the sequencing library

Synthesize the First Strand cDNA

1. Assemble the following reaction in a 200 μ L thin wall PCR tube:
Random Primers (2 μ L)

RNA + dH₂O (15µl)

The total volume should be 17µl.

2. Incubate the sample in a PCR thermal cycler at 65°C for 5 minutes, and then place the tube on ice.
3. Set the PCR thermal cycler to 25°C.
4. Add 1ul of SuperScript II to 9µl of 1st Strand Master Mix and mix well.
5. Add 8µl of pre-mixed SuperScript II and 1st Strand Master Mix into the 17µl primed RNA tube.
6. Incubate the sample in a thermal cycler with following program:
 - 25°C for 10 minutes
 - 42°C for 50 minutes
 - 70°C for 15 minutes
 - Hold at 4°C

Synthesize the Second Strand cDNA

1. Add 25 µl of thawed Second Strand Master Mix to the tube of 1st strand cDNA synthesis. Gently pipette the entire volume up and down 6 times to mix thoroughly.
2. Mix well and incubate at 16°C in a thermal cycler for 1 hour.

Purify double stranded cDNA

1. Vortex the AMPure XP beads until they are well dispersed and leave at room temperature at least 30 minutes.
2. Add 90 µl of well- mixed AMPure XP beads to each tube containing 50µl of double-stranded (ds) cDNA. Gently pipette the entire volume up and down 10 times to mix thoroughly.
3. Incubate the tube at room temperature for 15 minutes. Then transfer contents to a 1.5 ml microcentrifuge tube if necessary for convenience on the magnetic stand.
4. Place the tube on the magnetic stand at room temperature, for 5 minutes to make sure that all of the beads are bound to the side of the tube.
5. Remove and discard 135µl of the supernatant from the tube.
6. With the tube remaining on the magnetic stand, add 200µl of freshly prepared 80% ethanol to the tube without disturbing the beads.
7. Incubate the tube at room temperature for 30 seconds, then remove and discard all of the supernatant from each tube.
8. Repeat steps 5 and 6 once for a total of two 80% ethanol washes.
9. Leave the tube at room temperature for 15 minutes to dry and then remove the tube from the magnetic stand (Note: After removing all the liquid, do not let the beads

over-dry, as indicated when the pellet appears cracked. Elution efficiency will be significantly decreased at this point).

10. Centrifuge the thawed, room temperature Resuspension Buffer to 600 g for 5 seconds.
11. Add 52.5µl Resuspension Buffer to each tube and gently pipette the entire volume up and down 10 times to mix thoroughly.
12. Incubate the tube at room temperature for 2 minutes.
13. Place the tube on the magnetic stand at room temperature for 5 minutes.
14. Transfer 50µl of the supernatant (ds cDNA) from the tube to a new 0.2 ml PCR tube

SAFE STOPPING POINT: If you do not plan to proceed to next step immediately, the protocol can be safely stopped here. If you are stopping, store the sample at -20°C.

Perform End Repair

1. Add 10ul of Resuspension Buffer to each tube containing 50ul of ds cDNA
2. Add 40ul of End Repair Mix to tube containing the ds cDNA, mix thoroughly
3. Incubate on PCR machine at 30°C for 30 minutes.

Purify the End-Repaired cDNA

1. Vortex the AMPure XP beads until they are well dispersed and leave at room temperature at least 30 minutes.
2. Transfer end-repaired ds cDNA mixture from the 0.2 ml PCR tube to a 1.5 ml microcentrifuge tube. Add 160µl of well- mixed AMPure XP beads to each tube. Gently pipette the entire volume up and down 10 times to mix thoroughly.
3. Incubate the tube at room temperature for 15 minutes.
4. Place the tube on the magnetic stand at room temperature, for 5 minutes to make sure that all of the beads are bound to the side of the wells.
5. Using a 200µl pipette set to 127.5µl, remove and discard 127.5µl of the supernatant from each tube twice for total 255µl. (5ul remain)
6. With the tube remaining on the magnetic stand, add 200 µl of freshly prepared 80% ethanol to the tube without disturbing the beads.
7. Incubate the tube at room temperature for 30 seconds, then remove and discard all of the supernatant from each well.
8. Repeat steps 5 and 6 once for a total of two 80% ethanol washes.
9. Leave the tube at room temperature for 15 minutes to dry and then remove the tube from the magnetic stand.

10. Add 17.5µl room temperature Resuspension Buffer to each tube and gently pipette the entire volume up and down 10 times to mix thoroughly.
11. Incubate the tube at room temperature for 2 minutes.
12. Place the tube on the magnetic stand at room temperature for 5 minutes.
13. Transfer 15µl of the supernatant (end-repaired ds cDNA) from the tube to a new 0.2 ml PCR tube.

SAFE STOPPING POINT: If you do not plan to proceed to next step immediately, the protocol can be safely stopped here. If you are stopping, store the sample at -20°C .

Adenylate 3' Ends

1. Add 2.5µl of Resuspension Buffer to each tube containing end-repaired cDNA.
2. Add 12.5µl of A-Tailing Mix to the tube and mix thoroughly.
3. Incubate on PCR machine at 37°C for 30 minutes.
4. Immediately remove the tube from the thermal cycler, then proceed immediately to Ligate Adapters.

Ligate the Adapters

1. Add 2.5µl of Resuspension Buffer to each tube containing 3'-adenylated cDNA.
2. Add 2.5µl of Ligation Mix to each tube (only remove Ligation Mix tube from and put back to -15°C to -25°C immediately before and after each use).
3. Add 2.5µl of the appropriate RNA Adaptor Index to the tube and mix thoroughly.
4. Incubate on PCR machine at 30°C for 10 minutes.
5. Add 5µl of Stop Ligation Buffer to each tube and mix thoroughly.

Purify the Adapter-ligated cDNA

1. Vortex the AMPure XP beads until they are well dispersed and leave at room temperature at least 30 minutes.
2. Add 42µl of well-mixed AMPure XP beads to each tube containing 42.5µl of adapter-ligated ds cDNA. Gently pipette the entire volume up and down 10 times to mix thoroughly.
3. Incubate the tube at room temperature for 15 minutes. Then transfer contents to 1.5 ml microcentrifuge tube
4. Place the tube on the magnetic stand at room temperature, for 5 minutes to make sure that all of the beads are bound to the side of the wells.
5. Remove and discard 79.5µl of the supernatant from each tube.
6. With the tube remaining on the magnetic stand, add 200µl of freshly prepared 80% ethanol to the tube without disturbing the beads.

7. Incubate the tube at room temperature for 30 seconds, then remove and discard all of the supernatant from each well.
8. Repeat steps 5 and 6 once for a total of two 80% ethanol washes.
9. Leave the tube at room temperature for 15 minutes to dry and then remove the tube from the magnetic stand.
10. Add 50µl of room temperature Resuspension Buffer to each tube and gently pipette the entire volume up and down 10 times to mix thoroughly.
11. Incubate the tube at room temperature for 2 minutes.
12. Place the tube on the magnetic stand at room temperature for 5 minutes.
13. Transfer 20µl of the supernatant (adaptor-ligated cDNA) from the tube to the a 0.2 ml PCR tube

SAFE STOPPING POINT: If you do not plan to proceed to next step immediately, the protocol can be safely stopped here. If you are stopping, store the sample at –20°C.

Enrich the Purified cDNA Templates

1. Pre-program the thermal cycler with the following program:
 - Choose the pre-heat lid option and set to 100°C
 - 98°C for 30 seconds
 - 15 cycles of:
 - 98°C for 10 seconds
 - 60°C for 30 seconds
 - 72°C for 30 seconds
 - 72°C for 5 minutes
 - Hold at 10°C
2. Add 5µl of thawed PCR Primer Cocktail to each PCR tube.
3. Add 25µl of thawed PCR Master Mix to each tube. Gently pipette the entire volume up and down 10 times to mix thoroughly.
4. Place the PCR tube on the pre-programmed thermal cycler. Close the lid and select PCR to run the amplification cycle.

Clean up PCR

1. Vortex the AMPure XP beads until they are well dispersed and leave at room temperature at least 30 minutes.
2. Add 50µl of well- mixed AMPure XP beads to each tube containing 50µl of PCR amplified cDNA templates. Gently pipette the entire volume up and down 10 times to mix thoroughly.

3. Incubate the tube at room temperature for 15 minutes. Then transfer contents to 1.5 ml microcentrifuge tube.
4. Place the tube on the magnetic stand at room temperature, for 5 minutes to make sure that all of the beads are bound to the side of the wells.
5. Remove and discard 95µl of the supernatant from each tube.
6. With the tube remaining on the magnetic stand, add 200 µl of freshly prepared 80% ethanol to the tube without disturbing the beads.
7. Incubate the tube at room temperature for 30 seconds, then remove and discard all of the supernatant from each well.
8. Repeat steps 5 and 6 once for a total of two 80% ethanol washes.
9. Leave the tube at room temperature for 15 minutes to dry and then remove the tube from the magnetic stand.
10. Add 32.5µl of room temperature Resuspension Buffer to each tube and gently pipette the entire volume up and down 10 times to mix thoroughly.
11. Incubate the tube at room temperature for 2 minutes.
12. Place the tube on the magnetic stand at room temperature for 5 minutes.
13. Transfer 30µl of the clear supernatant (amplified library) to a new 1.5ml microcentrifuge tube.

SAFE STOPPING POINT: If you do not plan to proceed to next step immediately, the protocol can be safely stopped here. If you are stopping, store the sample at –20°C.

Validate the Library—Use a DNA specific chip such as the Agilent High Sensitivity DNA Chip on Agilent Technologies 2100 Bioanalyzer to validate the sequencing library as following:

Preparing the Gel-Dye Mix

1. Allow the blue-capped High Sensitivity DNA dye concentrate (blue) and red-capped High Sensitivity DNA gel matrix (red) to equilibrate to room temperature for 30 minutes in the dark.
2. Vortex the blue- capped vial with High Sensitivity DNA dye concentrate (blue) for 10 seconds and spin down. Make sure the DMSO is completely thawed.
3. Pipette 15 µl of the blue-capped dye concentrate (blue) into a red- capped High Sensitivity DNA gel matrix vial (red). Store the dye concentrate at 4 °C in the dark.
4. Cap the tube, vortex for 10 seconds. Visually inspect proper mixing of gel and dye.
5. Transfer the complete gel- dye mix to the top receptacle of a spin filter.
6. Place the spin filter in a microcentrifuge and spin for 10 minutes at room temperature at $2240\text{ g} \pm 20\%$ (for Eppendorf microcentrifuge, this corresponds to 6000 rpm).

7. Discard the filter according to good laboratory practices. Label the tube and include the date of preparation.

Loading the Gel-Dye Mix: Before loading the gel-dye mix, make sure that the base plate of the chip priming station is in position (C) and the adjustable clip is set to the lowest position.

1. Allow the gel- dye mix to equilibrate to room temperature for 30 minutes before use. Protect the gel- dye mix from light during this time.
2. Take a new High Sensitivity DNA chip out of its sealed bag and place the chip on the chip priming station.
3. Pipette 9.0 μ l of the gel- dye mix at the bottom of the well marked with the symbol 'G' and dispense the gel-dye mix.
4. Set the timer to 60 seconds, make sure that the plunger is positioned at 1 ml and then close the chip priming station. The lock of the latch will click when the Priming Station is closed correctly.
5. Press the plunger of the syringe down until it is held by the clip.
6. Wait for exactly 60 seconds and then release the plunger with the clip release mechanism.
7. Visually inspect that the plunger moves back at least to the 0.3 ml mark.
8. Wait for 5 s, then slowly pull back the plunger to the 1 mL position.
9. Open the chip priming station.
10. Pipette 9.0 μ L of the gel- dye mix in each of the wells marked with the symbol 'G'

Loading the Marker

1. Pipette 5 μ L of green- capped High Sensitivity DNA marker (green) into the well marked with the ladder symbol and into each of the 11 sample wells.

Loading the Ladder and the Samples

1. Pipette 1 μ l of the yellow- capped High Sensitivity DNA ladder vial (yellow) in the well marked with the ladder symbol.
2. In each of the 11 sample wells pipette 1 μ L of sample (used wells) or 1 μ L of marker (unused wells).
3. Place the chip horizontally in the adapter of the IKA vortex mixer and vortex for 60 seconds at 2400 rpm.
4. Run the chip in the Agilent 2100 Bioanalyzer. Make sure that the run is started within 5 minutes.

The typical final sequencing library on Bioanalyzer is shown in Figure 2.

BASIC PROTOCOL 3: DUPLEX-SPECIFIC THERMOSTABLE NUCLEASE (DSN) NORMALIZATION

The protocol here describes the use of DSN enzyme to reduce the high amount of rRNA present in the sequencing library. By increasing the temperature and then re-annealing the sequencing library, the high amount of cDNA molecules derived from rRNA in the library quickly form higher amounts of double stranded DNA than those from the less expressed host mRNA or those cDNAs derived from low amounts of viral RNA. Because DSN shows a strong preference for cleaving double strand DNA (Zhulidov, Bogdanova et al. 2004), applying DSN treatment in this step can greatly decrease the amount of cDNA derived from rRNA in the sequencing library.

Materials

Duplex-Specific Nuclease Kit (Axxora)
Agilent 2100 Bioanalyzer
Agilent High Sensitivity DNA Kit
AMPure XP beads
1M HEPES buffer solution
5M NaCl solution
Nuclease-free water
Heat block Magnetic stand
Thermal cycler
Phusion Polymerase
25 mM dNTPs
QIAGEN EB Buffer (10 mM Tris-Cl, pH 8.5)

DSN Treatment

1. Prepare the following 4xHybridization Buffer. Excess buffer can be prepared and stored for future use at -20°C .

Reagent	Volume (μl)
1 M HEPES buffer solution	200
5 M NaCl solution	400
Nuclease-free water	400
Total Volume Per Sample	1000

2. Ensure that the thermal cycler and heat block are located near each other.
3. Pre-heat the heat block to 68°C .

4. Prepare the following reaction mix in a separate, sterile, nuclease-free 200µl PCR tube on ice for each sample to be normalized.

Reagent	Volume (µL)
Sequencing library (80–100 ng)	13.5
4xHybridization Buffer	4.5
Total Volume Per Sample	18

5. Gently pipette the entire volume up and down 10 times, then centrifuge briefly to mix.
6. Transfer the entire volume of reaction mix directly to the bottom of a new, sterile, nuclease-free 200µl PCR tube, using a pipette. Do not let the sample contact the side of the tube during the process.
7. Incubate the reaction mix tube on the thermal cycler using the following PCR cycling conditions:
 - 98°C for 2 minutes
 - 68°C for 5 hours
8. Proceed immediately to following DSN treatment steps (CAUTION Following incubation, keep the thermal cycler lid closed and the temperature held at 68°C. Do not remove the reaction mix tube from thermal cycler prior to and during DSN treatment).
9. Dilute the 10X DSN Master buffer supplied in the DSN kit to 2X with nuclease-free water.
10. Pre-heat the 2X DSN buffer on the pre-heated heat block at 68°C.
11. Quickly add 20µl of pre-heated 2X DSN buffer to the first reaction mix tube.
12. With the reaction mix tube remaining within the thermal cycler, gently pipette the entire volume up and down 10 times to mix thoroughly using a pipette set to 40µl.
13. Incubate the reaction mix tube on the thermal cycler at 68°C for 10 minutes.
14. Quickly add 2µl of DSN enzyme to the reaction mix tube using a 2µl pipette.
15. With the reaction mix tube remaining within the thermal cycler, gently pipette the entire volume up and down 10 times to mix thoroughly using a pipette set to 40µl.
16. Incubate the reaction mix tubes on the thermal cycler at 68°C for 25 minutes.
17. Add 40µl of 2X DSN stop solution to each reaction mix tube. Gently pipette the entire volume up and down to mix thoroughly, then place the tubes on ice.

Purify DSN Treated DNA Templates

1. Vortex the AMPure XP beads until they are well dispersed and leave at room temperature at least 30 minutes.

2. Transfer contents of each reaction tube to 1.5 ml microcentrifuge tube if necessary.
3. Add 128 μ l of well- mixed AMPure XP beads to each tube containing 80 μ l of DSN treated library. Gently pipette the entire volume up and down 10 times to mix thoroughly.
4. Incubate the tube at room temperature for 5 minutes.
5. Place the tube on the magnetic stand at room temperature, for 5 minutes to make sure that all of the beads are bound to the side of the wells.
6. Remove and discard the supernatant from each tube. Take care not to disturb the beads.
7. With the tube remaining on the magnetic stand, add 180 μ l of freshly prepared 80% EtOH to the tube without disturbing the beads.
8. Incubate the tube at room temperature for 30 seconds, then remove and discard all of the supernatant from each well.
9. Repeat steps 5 and 6 once for a total of two 80% EtOH washes.
10. Let the tube at room temperature for 15 minutes to dry and then remove the tube from the magnetic stand.
11. Add 30 μ l of room temperature QIAGEN EB Buffer to each tube and gently pipette the entire volume up and down 10 times to mix thoroughly.
12. Incubate the tube at room temperature for 2 minutes.
13. Place the tube on the magnetic stand at room temperature for 5 minutes.
14. Transfer 30 μ l of the clear supernatant to the new microcentrifuge tubes

Enrich DNA Fragments

1. Prepare the following PCR reaction.

Reagent	Volume (μ L)
Cleaned DSN treated reaction mix	30
5X Phusion buffer	10
Phusion polymerase	0.5
25 mM dNTP	0.5
PCR Primer PE 1.0	0.5
PCR Primer PE 2.0	0.5
Nuclease-free water	8
Total Volume Per Sample	50

2. Amplify the PCR reaction on the thermal cycler using the following PCR cycling conditions
 - 98°C for 30 seconds

- 12 cycles of:
 - 98°C for 10 seconds
 - 60°C for 30 seconds
 - 72°C for 30 seconds
- 72°C for 5 minutes
- Hold at 10°C

Purify PCR Products

1. Vortex the AMPure XP beads until they are well dispersed and leave at room temperature at least 30 minutes.
2. Transfer contents of each reaction tube to 1.5 ml microcentrifuge tube if necessary.
Transferring to the 1.5 ml tube depends on the size of your magnetic stand. It is more convenient to use a 1.5 ml tube if the stand is designed to fit those tubes.
3. Add 80 µl of well-mixed AMPure XP beads to each tube containing 50 µl of PCR amplified library. Gently pipette the entire volume up and down 10 times to mix thoroughly.
4. Incubate the tube at room temperature for 5 minutes.
5. Place the tube on the magnetic stand at room temperature, for 5 minutes to make sure that all of the beads are bound to the side of the wells.
6. Remove and discard the supernatant from each tube. Take care not to disturb the beads.
7. With the tube remaining on the magnetic stand, add 180 µl of freshly prepared 80% EtOH to the tube without disturbing the beads.
8. Incubate the tube at room temperature for 30 seconds, then remove and discard all of the supernatant from each well.
9. Repeat steps 5 and 6 once for a total of two 80% EtOH washes.
10. Leave the tube at room temperature for 15 minutes to dry and then remove the tube from the magnetic stand.
11. Add 20 µl of room temperature QIAGEN EB Buffer to each tube and gently pipette the entire volume up and down 10 times to mix thoroughly.
12. Incubate the tube at room temperature for 2 minutes.
13. Place the tube on the magnetic stand at room temperature for 5 minutes.
14. Transfer 30 µl of the clear supernatant to the new 1.5 ml microcentrifuge tubes. Store at –20°C.
15. Repeat duplex-specific thermostable nuclease (DSN) normalization one more time by repeating all the steps in BASIC PROTOCOL 3).

16. Validate the final sequencing library on an Agilent 2100 Bioanalyzer using the Agilent high sensitivity DNA chip by repeating all the steps in Validate the Library section in BASIC PROTOCOL 2.
17. Cluster the final library on the cbot and sequence on an Illumina high throughput sequencer following standard Illumina protocols (<http://www.illumina.com/>).

REAGENTS AND SOLUTIONS

Extraction buffer (20 mM Tris-HCL pH 7.6, 20 mM EDTA, 1% SDS)

To make 100 ml:

- 2 ml 1M Tris, pH 7.6
- 4 ml 0.5 M EDTA
- 10 ml 10% SDS
- 84 ml Molecular-grade water
- Extraction buffer is stored in 1.5 ml aliquots at room temperature without Proteinase K. Before use Proteinase K stocks are thawed and 45µl of Proteinase K stock (20 mg/ml) is added to 1.5 ml of extraction buffer.

COMMENTARY

Background Information

FFPE advantages and disadvantages—Formalin-fixed, paraffin-embedded (FFPE) specimens provide valuable archival materials that may go back as far as the late 19th century. It is estimated that, worldwide, over a billion tissue samples, most of them FFPE, are being stored in numerous hospitals, tissue banks, and research laboratories. These archived samples could potentially provide a wealth of information in retrospective molecular studies of different diseases. FFPE tissue is relatively stable for decades, is easily handled, long-term storage is inexpensive, and it is suitable for immunohistochemical or molecular analyses with a low cost of large-scale applications (Kayser, Stute et al. 1988; Perlmutter, Best et al. 2004). In addition, most FFPE samples have associated pathological and clinical annotations, which is very important for retrospectively studying these diseases. However, the big challenges of working with FFPE samples are the cross-linking of nucleic acids due to formalin fixation and severe RNA degradation during the embedding process (Foss, Guha-Thakurta et al. 1994; McKinney, Moon et al. 2009). In addition, formalin fixation modifies RNA by adding methylol groups that can interfere with reverse transcription (Masuda, Ohnishi et al. 1999).

High-throughput sequencing technology, which can determine millions of bases (megabases) of DNA sequence per run (Service 2006), has been applied mostly to cancer-related FFPE samples to perform DNA-seq (Schweiger, Kerick et al. 2009; Wood, Belvedere et al. 2010; Singh, Patel et al. 2013) and miRNA-seq (Weng, Wu et al. 2010). However, characterization and recovery of full length of microbial genomes (e.g., viral and bacterial) from FFPE sample by high throughput sequencing is more challenging. In this protocol, we describe methods that allowed the complete genome of the 1918 pandemic

influenza virus from an archival FFPE specimen to be determined using a high-throughput sequencing approach (Xiao, Kash et al. 2013). The RNA isolated from such archival 1918 FFPE sample of 1918 “Spanish flu” pandemic is highly degraded (Taubenberger, Reid et al. 1997). From our experience, the length profile of isolated RNAs from these samples is around 100 nucleotides or less. In addition, the RNA isolated from such FFPE tissue samples are predominantly ribosomal RNA and tRNA (Tariq, Kim et al. 2011). Direct sequencing of a total cDNA library will generate mainly rRNA reads with few influenza viral RNA reads. The standard mRNA-seq procedure using oligo dT primers for mRNA poly-A selection is also not applicable because the RNA isolated from FFPE is highly degraded, leading to production of an extremely 3’ biased sequence library with little or no ability to recover full length viral genome sequences.

Therefore, we applied Duplex-Specific Thernostable Nuclease (DSN) (Evrogen, Moscow, Russia) (Zhulidov, Bogdanova et al. 2004) normalization procedures on the sequencing library to reduce the amount of rRNAs present in the sample. DSN has been used to normalize cDNA to enhance the detection of rare transcripts in eukaryotic cDNA libraries by decreasing the prevalence of highly abundant transcripts (Zhulidov, Bogdanova et al. 2004). This DSN method includes the denaturation of cDNA, its subsequent reassociation and enzymatic degradation of the double-stranded DNA (dsDNA) fraction using DSN isolated from the Kamchatka crab (Shagin, Rebrikov et al. 2002). Because abundant transcripts form dsDNA more effectively during the reassociation step, they are easily subjected to DSN-mediated degradation. It has been shown that DSN treatment is a better method to remove rRNA than conventional subtractive hybridization methods for construction of high-throughput sequencing libraries in prokaryotic systems (Yi, Cho et al. 2011). However, a recent study showed that Ribo-Zero (Epicentre) rRNA removal method outperformed a DSN normalization method in rRNA depletion efficiency using FFPE samples (Zhao, He et al. 2014), which is a promising alternative. However, the reliability and repeatability of these different protocols need validation in each lab attempting enrichment strategies.

Critical Parameters and Troubleshooting

RNA input amount—The quality and quantity of RNA isolated from archival FFPE samples is highly dependent on the archival source material, with uncontrollable factors including type of tissue, time spent prior to fixation, type and length of fixation, storage conditions, and age of the block (Krafft, Duncan et al. 1997). From our recent experience constructing cDNA libraries for high-throughput sequencing, newer FFPE samples, for example autopsy materials from the 2009 influenza pandemic, are much easier to work with than very old samples like autopsy FFPE samples from the 1918 influenza pandemic. The amount of RNA used for direct high-throughput library construction that we used is between 100ng to 300ng, which are typical amounts of RNA that we can obtain from FFPE materials. Using this amount of input RNA, we usually can construct satisfactory cDNA libraries for high-throughput sequencing.

Final cDNA library amount and PCR amplification—Another critical point in the process is the final cDNA library amount, which is critical for subsequent DSN treatment.

From our experience, the final library amount needs to be in the range of 50ng to 800ng. Infrequently, the amount of final cDNA library is less than 50ng. In these cases, one could either re-make the library from the beginning by increasing the amount of input RNA or alternatively increase the number of PCR cycles used to achieve enhanced amplification. The highest number of cycles of the final PCR step before DSN treatment that we have used is 18 cycles. The amount of input DNA required for DSN treatment is 80–100ng in 13.5µl. Therefore, if the library is only 50ng, vacuum concentration can be used to concentrate the library from 30 µl to about 15 µl final volume. The DNA amount is critical for DSN treatment because the hybridization rate for each nucleotide molecule is proportional to the square of its concentration (Young and Anderson 1985).

DSN treatment—DSN treatment was performed twice in our experiments (Xiao et al., 2013). However, the number of DSN treatments is dependent upon the following factors: 1) how much rRNA is in the library; 2) how rare the viral RNA is in the library; and 3) how many reads will be generated. Use of a quantitative real-time PCR assay is suggested to check the results of DSN treatment of the library. The comparison C_T levels of a single copy gene (for example, human β 2 microglobulin) (Krafft, Duncan et al. 1997) and human 18s rRNA (Xiao, Kash et al. 2013) can be utilized to monitor the effectiveness of DSN treatment.

Anticipated Results

The protocols described here are very practical and useful for isolating RNA from various FFPE samples including those as old as ~100 years old and constructing rRNA-reduced total RNA-derived cDNA high throughput sequencing library for the Illumina platform. Normally, one to three 6-micron sections of a FFPE sample are used to isolate RNA and 200ng to 700ng total RNA can be produced. From 100ng to 300ng input RNA, a sequencing library can be successfully made with a concentration of at least ~10ng/µl and ~50nM, which is enough to perform the cluster generation step. However, the number of reads of the viral genome sequences will vary depending on the archival samples utilized and the total number of reads generated by the sequencer.

Time Considerations

The whole protocol, from isolation of RNA to finishing DSN treatment of sequencing library, takes 4 or 5 days to complete depending on DSN treatment once or twice. Basic Protocol 1: isolation of RNA from archival fixed tissue samples, takes 2 days to finish. Basic Protocol 2: sequencing library preparation, takes 2 days to finish. Basic Protocol 3: duplex-specific thermostable nuclease (DSN) normalization (once), takes 1 day to finish. In addition, during the whole protocol, there are various safe stopping points, at which times sample can be stored at -20°C .

Acknowledgments

This work was supported by the Intramural Research Program of the National Institute of Allergy and Infectious Diseases, National Institutes of Health.

LITERATURE CITED

- April C, Klotzle B, et al. Whole-genome gene expression profiling of formalin-fixed, paraffin-embedded tissue samples. *PLoS ONE*. 2009; 4(12):e8162. [PubMed: 19997620]
- Baldwin DA, Feldman M, et al. Metagenomic assay for identification of microbial pathogens in tumor tissues. *mBio*. 2014; 5(5):e01714–01714. [PubMed: 25227467]
- Beck AH, Weng Z, et al. 3'-end sequencing for expression quantification (3SEQ) from archival tumor samples. *PLoS ONE*. 2010; 5(1):e8768. [PubMed: 20098735]
- Bibby K. Metagenomic identification of viral pathogens. *Trends in biotechnology*. 2013; 31(5):275–279. [PubMed: 23415279]
- Evers DL, Fowler CB, et al. The effect of formaldehyde fixation on RNA: optimization of formaldehyde adduct removal. *The Journal of molecular diagnostics : JMD*. 2011; 13(3):282–288. [PubMed: 21497290]
- Evers DL, He J, et al. Paraffin embedding contributes to RNA aggregation, reduced RNA yield, and low RNA quality. *The Journal of molecular diagnostics : JMD*. 2011; 13(6):687–694. [PubMed: 21884819]
- Firth C, Lipkin WI. The genomics of emerging pathogens. *Annual review of genomics and human genetics*. 2013; 14:281–300.
- Foss RD, Guha-Thakurta N, et al. Effects of fixative and fixation time on the extraction and polymerase chain reaction amplification of RNA from paraffin-embedded tissue. Comparison of two housekeeping gene mRNA controls. *Diagnostic molecular pathology : the American journal of surgical pathology, part B*. 1994; 3(3):148–155.
- Fox CH, Johnson FB, et al. Formaldehyde fixation. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society*. 1985; 33(8):845–853. [PubMed: 3894502]
- Kayser K, Stute H, et al. Rapid microwave fixation--a comparative morphometric study. *The Histochemical journal*. 1988; 20(6–7):347–352. [PubMed: 3065304]
- Krafft AE, Duncan BW, et al. Optimization of the Isolation and Amplification of RNA From Formalin-fixed, Paraffin-embedded Tissue: The Armed Forces Institute of Pathology Experience and Literature Review. *Mol Diagn*. 1997; 2(3):217–230. [PubMed: 10462613]
- Masuda N, Ohnishi T, et al. Analysis of chemical modification of RNA from formalin-fixed samples and optimization of molecular biology applications for such samples. *Nucleic Acids Res*. 1999; 27(22):4436–4443. [PubMed: 10536153]
- McKinney MD, Moon SJ, et al. Detection of viral RNA from paraffin-embedded tissues after prolonged formalin fixation. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology*. 2009; 44(1):39–42. [PubMed: 18977691]
- Mittempergher L, de Ronde JJ, et al. Gene expression profiles from formalin fixed paraffin embedded breast cancer tissue are largely comparable to fresh frozen matched tissue. *PLoS ONE*. 2011; 6(2):e17163. [PubMed: 21347257]
- O'Leary, TJ. Infectious Diseases. In: O'Leary, TJ., editor. *Advanced Diagnostic Methods in Pathology*. Philadelphia: Saunders; 2003. p. 159-190.
- Perlmutter MA, Best CJ, et al. Comparison of snap freezing versus ethanol fixation for gene expression profiling of tissue specimens. *The Journal of molecular diagnostics : JMD*. 2004; 6(4):371–377. [PubMed: 15507677]
- Schweiger MR, Kerick M, et al. Genome-wide massively parallel sequencing of formaldehyde fixed-paraffin embedded (FFPE) tumor tissues for copy-number- and mutation-analysis. *PLoS ONE*. 2009; 4(5):e5548. [PubMed: 19440246]
- Service RF. Gene sequencing. The race for the \$1000 genome. *Science*. 2006; 311(5767):1544–1546. [PubMed: 16543431]
- Shagin DA, Rebrikov DV, et al. A novel method for SNP detection using a new duplex-specific nuclease from crab hepatopancreas. *Genome research*. 2002; 12(12):1935–1942. [PubMed: 12466298]
- Singh RR, Patel KP, et al. Clinical validation of a next-generation sequencing screen for mutational hotspots in 46 cancer-related genes. *The Journal of molecular diagnostics : JMD*. 2013; 15(5):607–622. [PubMed: 23810757]

- Tang W, David FB, et al. DNA extraction from formalin-fixed, paraffin-embedded tissue. Cold Spring Harbor protocols. 2009; 2009(2) pdb prot5138.
- Tariq MA, Kim HJ, et al. Whole-transcriptome RNAseq analysis from minute amount of total RNA. Nucleic Acids Res. 2011; 39(18):e120. [PubMed: 21737426]
- Taubenberger JK, Reid AH, et al. Initial genetic characterization of the 1918 “Spanish” influenza virus. Science. 1997; 275(5307):1793–1796. [PubMed: 9065404]
- Weng L, Wu X, et al. MicroRNA profiling of clear cell renal cell carcinoma by whole-genome small RNA deep sequencing of paired frozen and formalin-fixed, paraffin-embedded tissue specimens. J Pathol. 2010; 222(1):41–51. [PubMed: 20593407]
- Wood HM, Belvedere O, et al. Using next-generation sequencing for high resolution multiplex analysis of copy number variation from nanogram quantities of DNA from formalin-fixed paraffin-embedded specimens. Nucleic Acids Res. 2010; 38(14):e151. [PubMed: 20525786]
- Xiao YL, Kash JC, et al. High-throughput RNA sequencing of a formalin-fixed, paraffin-embedded autopsy lung tissue sample from the 1918 influenza pandemic. J Pathol. 2013; 229(4):535–545. [PubMed: 23180419]
- Yi H, Cho YJ, et al. Duplex-specific nuclease efficiently removes rRNA for prokaryotic RNA-seq. Nucleic Acids Res. 2011; 39(20):e140. [PubMed: 21880599]
- Young, BD.; Anderson, M. Quantitative analysis of solution hybridisation. In: Hames, BD.; Higgins, SJ., editors. Nucleic Acids Hybridisation, a Practical Approach. Oxford: IRL Press; 1985. p. 47-71.
- Zhao W, He X, et al. Comparison of RNA-Seq by poly (A) capture, ribosomal RNA depletion, and DNA microarray for expression profiling. BMC Genomics. 2014; 15:419. [PubMed: 24888378]
- Zhulidov PA, Bogdanova EA, et al. Simple cDNA normalization using kamchatka crab duplex-specific nuclease. Nucleic Acids Res. 2004; 32(3):e37. [PubMed: 14973331]

Figure 1.
Typical FFPE total RNA on Bioanalyzer (The small peak in front is the RNA lower marker).