WPI

Direct RNA Sequencing and Transcriptomic Analysis of *Pseudomonas putida*

A Major Qualifying Project submitted to the faculty of
Worcester Polytechnic Institute in partial fulfillment of the
requirements for the degree of Bachelor of Science

Submitted By:
Maia Grant – Bioinformatics & Computational Biology
Stokley Voltmer – Bioinformatics & Computational Biology

Approved By:
Natalie Farny, PhD. – Department of Biology & Biotechnology
Eric Young, PhD. – Department of Chemical Engineering

Date Submitted:
April 29, 2022

Abstract

The transcriptome of a cell is the total messenger RNA (mRNA) content of the cell. The transcriptome is useful to understand the changes that can occur in gene expression, specifically amongst growth conditions. Gene expression can be analyzed through many sequencing techniques. One technique is direct RNA sequencing (direct RNA-seq), a relatively new technique that does not need RNA to be converted to cDNA and does not rely on amplification methods of any kind. Direct RNA-seq has not been utilized to perform analysis on the transcriptome in the soil environment. Typically, this method is not the preferred method for bacterial RNA sequencing, mostly due to the need for rRNA depletion, lowering the concentration of pure RNA samples by at least 90%. Direct RNA-seq was attempted to investigate the transcriptome of Pseudomonas putida in various growth conditions. This project was completed in hopes that the transcriptomic data acquired would be used to engineer improved biosensors for environmental applications. After several attempts at isolating viable RNA for sequencing, it was determined that the concentration of the samples acquired was too low for rRNA depletion, making it impossible to sequence. Instead, we analyzed previously acquired transcriptomic data of species native to the soil environment to find any similarities amongst them. Our paper presents the attempted methods used to isolate RNA, along with our recommendations for further projects to consider.

Acknowledgements

We would like to thank the following people who made this project possible. Professor Natalie Farny for her guidance and support throughout the duration of the project. Professor Eric Young for his insight and knowledge of direct RNA-sequencing and his guidance on the Oxford Nanopore minION system and the KBase platform. Finally, Matthew Bruno for assisting us using his knowledge of sample purification, demonstrating various adapted methods from his own research, and answering our many questions on where to find different materials in the lab.

Table of Contents

Abstract **II**

Acknowledgements **III**

Table of Contents **1**

Introduction **3**

Transcriptome 3

Methods of rRNA Depletion 3

Direct RNA Sequencing 4

Previous Related Studies 4

Purpose of the Present Study 5

Methodology **6**

Solubilized Extract of Soil Organic Material (SESOM) 6

Bacteria Sample Growth 6

LB Growth 6

SESOM Growth 7

Cell Lysis 7

RNA Purification from Bacterial Lysate 7

gDNA Removal and DNase Treatment 7

PCR 8

Gel Electrophoresis 9

RNA Cleanup 9

Polyadenylation 10

KBase and Analysis 10

Isolation of Pure RNA 12

PCR Analysis 13

Polyadenylation 14

Sample Concentration 15

Discussion **16**

References **19**

Appendices **22**

Appendix A: Direct RNA- Sequencing vs Illumina Sequencing 22

Appendix B: Previous Relevant Studies Using Direct RNA-Sequencing on Bacterial Species 24

Appendix C: Flowchart of Methods Used in This Project 26

Appendix D: Protocol with Notes for Possible Modifications 27

Appendix E: PCR Primer Set for P. putida KT2440 Gene PP_2088 33

Appendix F: Transcriptomic Analysis 34

Appendix G: Protocol for Phenol: Chloroform RNA Extraction 36

Appendix H: Estimation of DNA and RNA in Samples 38

Introduction

Pseudomonas putida, a fast-growing bacterium with low nutrient demand commonly found in both soil and water, is a candidate host for many biotechnological uses (Poblete-Castro et al., 2017). Current research in the Young and Farny laboratories at WPI, regarding *P. putida*, focuses on the use of *P. putida* as a chemical sensor. *P. putida* KT2440 has been used for producing target chemicals due to its high toxicity tolerance and high natural flux in aromatic-catabolic pathways (Martínez-García and de Lorenzo, 2017, Nikel et al., 2016). Previous work completed by now graduated lab member Kyle Harding compared existing RNA-seq data from across a variety of growth conditions for the purpose of identifying constitutively expressed promoters (Harding, 2021). This project zoomed out from the work previously done and will address the difference in transcriptome between *P. putida* grown in lab conditions and *P. putida* grown in soil, its environmental setting. The method of using direct RNA sequencing has not been previously utilized to analyze the *P. putida* transcriptome when grown in soil.

Microbes observed in the laboratory may not exhibit growth and behavior as encoded in the genome. Along with this, some genomic features can only be expressed in situ. Through transcriptomics, we are attempting to show functions which are specifically expressed in an environmental setting rather than in the laboratory (Kalyuzhnaya et al., 2009). Gene expression in situ is a direct approach for determining the function of *P. putida*, which will allow us to build synthetic gene circuits that function optimally in soil. The growth condition differences will ultimately allow us to know how to better engineer biosensor technologies.

Transcriptome

The transcriptome of a cell is the total messenger RNA (mRNA) content of the cell. Because the first step of gene expression is transcribing the genes to mRNA, the transcriptome is an indication of gene expression. Unlike proteins, mRNA degrades rapidly, meaning that the transcriptome is sensitive to changes in gene expression.

Transcriptomic analysis will allow us to determine how the gene expression of *P. putida* changes in response to the soil environment. Challenges of transcriptomics in a soil organism include the incomplete knowledge that exists with their genome sequences, especially in their natural environment (Shakya et al., 2019). In order to have fewer reads when sequencing RNA of *P. putida*, ribosomal RNA (rRNA) will need to be depleted.

Methods of rRNA Depletion

One difficulty inherent to RNA sequencing is the presence of ribosomal RNA. rRNA, which composes the core of the ribosome complex and catalyzes protein synthesis, compromises upward of 90% of RNA-seq reads (Sarode et al., 2016) and provides no insight into changes in

gene expression. The significant percentage of rRNA limits the number of mRNA reads and produces high total read depth. Because of these complications, it is desirable to remove the rRNA prior to performing the sequencing. To this end, commercial kits have been designed to deplete rRNA in bacteria. Additionally, several low-cost methods have been published for use with bacteria, most commonly strains of E. coli (Wangsanuwat et al., 2020). This project intended to use the commercial NEBNext rRNA depletion kit, which uses the NEBNext RNase H-based RNA depletion workflow to deplete 5S, 16S, and 23S rRNA (Biolabs, New England, 2021).

Direct RNA Sequencing

Direct RNA sequencing will allow us to analyze the transcriptome of *P. putida*. This method allows for full-length transcripts (Oxford Nanopore, 2021). Along with this, the reads will be accessed fast due to real-time analysis, which reads at about 400 bases/second (Oxford Nanopore, 2022). For more information regarding direct RNA sequencing and how it differs from Illumina sequencing, refer to Appendix A.

Comprised of hundreds of nanopores contained within a flow cell, the minION platform is an efficient method of DNA and RNA sequencing. Through the course of sequencing a DNA or RNA strand passes through a nanopore, and the minION records the change in ionic current at each instant. A base-calling algorithm decodes the changes in ionic current and returns the resulting decoded sequence to the user (Jain et al., 2016).

Prior to sequencing, any libraries of interest must be prepared. The most important part of this preparation is the ligation of sequence adapters to the sequences of interest. The sequence adapters act both as the anchor point that the nanopore attaches to, and the molecular engine that pushes the strand through the nanopore (Jain et al., 2016). Currently, the poly(A) tail of mRNAs is used as an anchor for direct sequencing on the Nanopore platform. Prokaryotic mRNAs are not polyadenylated, which presents an additional challenge in using direct RNA sequencing with *P. putida*.

Previous Related Studies

Few studies currently exist on using direct RNA-seq of bacterial species with the Nanopore series. Even fewer focus on studying the transcriptome itself. In a study conducted by researchers at the University of Regensburg, the team presented findings in hopes to implement the direct RNA-seq technique for prokaryotic transcriptomes (Grünberger et al., 2022). Another study performed by researchers in Australia studied the effectiveness of bacterial direct RNA sequencing and base-calling to evaluate the genome and resistome of an extensively drug-resistant bacterial strain (Pitt et al., 2020). Lastly, a study conducted at the UC Santa Cruz

Genomics Institute utilized direct MinION nanopore sequencing of individual, full-length 16S rRNA to overcome epigenetic modifications on rRNA that are erased by sequencing-by-synthesis methods (Smith et al., 2019). These studies were successfully able to generate pure bacterial RNA samples that were able to be sequenced. To date, a study has not been published working with *P. putida*, or any other soil microbe for that matter. For further information and the major findings of these and other previous studies, along with a shortened version of their methods, refer to Appendix B.

Purpose of the Present Study

The gene expression of *P. putida* was intended to be determined by means of transcriptome analysis. The transcriptome was to be quantified by extracting and purifying mRNA from *P. putida* and sequencing the samples using the Nanopore minION system. We expected that gene expression, observable through the transcriptome, would vary through different soil and laboratory growth conditions. We planned on analyzing the expression using a pipeline hosted through the KBase service. Upon the failure to retrieve a highly concentrated RNA sample in either growth condition, the project shifted to using the pipeline to analyze the known transcriptomic sequences of bacterial species naturally found in the soil environment. The goal of the analysis was to validate the work done previously by Harding and to find a gene that was highly expressed under many growth conditions with low variability in expression between conditions. A gene that meets said requirements may be well suited to use in genetic engineering applications in the soil environment.

Methodology

The methods listed below were adapted over the process of the project. For the order of the methods followed, refer to Appendix C for the process of retrieving samples for direct RNA-seq. For a compiled list of the protocols followed, including suggested steps that could be attempted, along with the protocol for rRNA depletion, refer to Appendix D. The protocols used for RNA extraction are adapted from a study that extracted the RNA from *E. coli*, for use in RT-qPCR. For validation purposes, ~ 3 µL of each method was stored in a centrifuge tube for later use.

Solubilized Extract of Soil Organic Material (SESOM)

Solubilized Extract of Soil Organic Material, or SESOM, was made in the lab using store-bought soil. This was produced by adding 100 g of soil with 500 mL of 10X phosphate-buffered saline (PBS) buffer to an Erlenmeyer flask and placed in a 37ºC incubated shaker for 2-4 hours. This mixture was then added to a French press to separate the liquid, decanting slowly to ensure large solids stayed behind. Cytiva Whatman circular qualitative filter paper Grade 1 (particle retention <11µm) was placed under a piece of Grade 4 circular qualitative filter paper (particle retention <24µm) in a Buchner funnel. The funnel was placed on a rubber stopper to adhere to a filtration flask. The filter paper was then saturated around the edges with PBS to ensure they were sealed. The solution was then vacuum filtered overnight. The remaining liquid is SESOM, which was stored in a 4ºC fridge for later use in culture growth.

Bacteria Sample Growth

LB Growth

To begin, 10 mL culture tubes were filled with 5 mL of LB media, and the tubes were inoculated with cultures of *P. putida* from the glycerol stock. The cultures were incubated at 37ºC with shaking overnight. After culture growth, the culture was back diluted. To do this, 0.5 mL of the culture was removed from the culture tube, placed into a fresh 10 mL culture tube containing 5 mL of LB media, and placed in a 37ºC shaker for 90 minutes. After 90 minutes, the absorbance was measured in a spectrophotometer at 260 nm. Once the optical density (OD) reached about 0.5 (mid-log phase), 1 mL samples were placed into the spectrophotometer, the OD was recorded, and the samples were placed into 2 mL centrifuge tubes. Cell pellets were then created from each sample. To do this, each tube was placed into the centrifuge and centrifuged at 10,000 rpm for 10 minutes. The supernatant was decanted and discarded, ensuring not to disturb the pellet. The pellets were stored at -80ºC until further use.

SESOM Growth

To begin, 10 mL culture tubes were filled with 5 mL of SESOM media, the tubes were inoculated with cultures of *P. putida* from the glycerol stock and incubated at 37ºC with shaking overnight. After culture growth, 1 mL of the culture was removed, and the OD was measured at 260 nm. It is noted that this culture did not need to go through the back dilution process, due to the slow growth of *P. putida* in SESOM. Once the optical density (OD) reached about 0.5, 1 mL samples were placed into the spectrophotometer, the OD was recorded, and samples were placed into 2 mL centrifuge tubes. Cell pellets were created from each sample. To do this, each tube was placed into the centrifuge and centrifuged at 10,000 rpm for 10 minutes. The supernatant was decanted and discarded, ensuring not to disturb the pellet. The pellets were stored at -80ºC until further use.

Cell Lysis

Bacterial lysis of *P. putida* was performed to remove total RNA in the samples. RNaseZap was used on lab bench, pipettes, and gloves to prevent RNase contamination. First, 200 µL of lysozyme-TE buffer was added to centrifuge tubes containing the pellet. Next, 10µL of proteinase K was added to each tube. Samples were pipetted to resuspend the pellet, then mixed by vortexing for 10 seconds and incubated at room temperature for 10 minutes. During incubation, the centrifuge tubes were vortexed for 10 seconds every 2 minutes. Lastly, 700 µL of Buffer RLT and 7 µL of β-mercaptoethanol were added to the tubes to stop the reaction and were then vortexed again.* The lysate was then placed into a genomic DNA (gDNA) eliminator spin column and centrifuged for 30 seconds at 8,000 x g. The supernatant was then treated with DNase.

* The adapted protocol states that Buffer RLT is only stable for 1 month after the addition of β-mercaptoethanol to the samples.

RNA Purification from Bacterial Lysate

gDNA Removal and DNase Treatment

The DNase treatment was adapted from manufacturer's instructions from the TURBO DNA-free kit and modified for use with *P. putida*. It is noted that better results were displayed when the DNase treatment was performed before the gDNA removal. If the nucleic acid solution concentration was >200 µg/mL, the sample was diluted to 10 ng nucleic acid/50 µL. The DNase treatment can be performed on 20 µL of extract at a time. After this, 2 µL of 10X TURBO DNase Buffer was added (to 1X concentration) in the RNA sample. Next, 1 µL of TURBO DNase was added to the RNA. The sample was mixed by inverting and incubated for 25 minutes at 37ºC in a

dry heat bath. To inactivate the TURBO DNase, a 15mM concentration of EDTA was added. First, 0.72 µL of EDTA was added, then 0.28 µL of RNase- free water. This was incubated in a dry heat bath set at 75ºC for 10 minutes. Samples then went through a contamination removal process.

To begin the gDNA removal process, the concentration of each RNA sample was calculated using the Thermo Scientific™ NanoDrop™ One Microvolume UV-Vis Spectrophotometer by first blanking the reader with PBS and then placing 2 mL of lysate onto the reader. The sample was diluted to get 50 mL of a sample with a concentration of 100ng/µL. After this, 500 µL of 98% ethanol was added to the tube and mixed by pipetting. Next, 700 µL of diluted lysate containing ethanol was transferred into a RNeasy Mini spin column and centrifuged for 15 seconds at 8000 x g. The flow-through was discarded. Next, 700 µL of Buffer RW1 was added to the RNeasy spin column using the same collection tube. After incubating at room temperature for 5 min, the tubes were centrifuged for 15 seconds at 8000 x g. The collection tube containing the flow-through was discarded. The column was placed in a new collection tube. Next, 500 µL of Buffer RPE with ethanol was added to the column and centrifuged for 15 seconds at 8000 x g and the flow-through was discarded. Next, 500 µL Buffer RPE was added again to the column. The tube containing the column was then centrifuged for 2 min at 8000 x g. According to the manufacturer's troubleshooting guide, the column should be spun one more time to rid any excess, unwanted liquid from the column. For this, the column was placed in a fresh tube and was centrifuged again for 1 minute at 8000 x g. The column was placed into a new 1.5 mL RNase-free collection tube and 30 µL of RNase-free water was added directly onto the column and sat in room temperature for 10 minutes. After this, the tube was centrifuged for 1 min at 8000 x g. The flow-through was collected and placed on the column again, in an attempt to increase the sample concentration, and the sample sat at room temperature for 10 minutes. After this, the tube was centrifuged for 1 min at 8000 x g.

PCR

This protocol was adapted from One*Taq* Quick-Load 2X Master Mix with Standard Buffer from New England BioLabs with adjustments. Reaction components were kept on ice. First, to a 0.2 mL PCR tube, 12.5 µL of One*Taq* Quick-Load 2X Master Mix with Standard Buffer was added, along with 0.5µL (10µM concentration) of forward primer, 0.5µL (10µM concentration) of reverse primer, 1µl of template DNA, and 10.5 µl of Nuclease-free water. A positive control sample containing cDNA created by Bruno, and a negative control sample were used as well. For a negative control, 12.5µL of One*Taq* Quick-Load 2X Master Mix with Standard Buffer was added, along with 0.5µL (10µM concentration) of forward primer, 0.5µL (10µM concentration) of reverse primer, and 11.5 µl of Nuclease-free water were added to a 0.2

mL PCR tube. The primer set is for gene PP_2088 (sequence listed in Appendix E). Each reaction was gently mixed by pipetting. The sample was placed in a thermocycler with the following conditions as listed in Table 1:

Step	Temperature	Time
Initial Denaturation	94ºC	30 seconds
30 Cycles	94ºC 67-68ºC 68ºC	20 seconds 20 seconds 15 seconds
Final Extension	68ºC	5 minutes
Hold	4ºC	∞

Table 1: Thermocycling Conditions for PCR. This table shows the thermocycling conditions for a routine PCR, which are also used with OneTaq Quick-Load 2X Master Mix with Standard Buffer from New England BioLabs.

Gel Electrophoresis

To make the agarose gel for gel electrophoresis, an Erlenmeyer flask was placed on an electronic balance, and the scale was zeroed. For one 1% agarose gel, 0.5 g of agarose was measured. To the agarose, 50 mL 1XTAE was added. The scale was zeroed again. After this, the flask was microwaved for 1 minute, swirled to mix, and microwaved again for 1 min. The flask was swirled again and microwaved for 20 seconds more. Next, the flask was placed on the scale and deionized water was added until the balance read zero. After this, 2.5 µL of SYBR Safe was added to the agarose mixture and swirled. The agarose mixture was poured into a gel tray with a well comb in place. The gel sat at room temperature for 20-30 minutes. After the gel had set, it was placed in the chamber. The electrophoresis chamber was filled with 1XTAE Buffer until the gel was covered. The samples were loaded into the wells (5 µL for PCR samples and 10 µL for samples we used for confirmation of gDNA removal), including 2 µL of the GeneRuler 1kb DNA Ladder. Each gel was run at 100 V for 30 minutes. The gel was imaged using a Bio-Rad GelDoc machine using UV light and analyzed for gDNA contamination. For gel images, refer to the Results section.

RNA Cleanup

The RNeasy Mini Kit was used to clean up RNA previously isolated through the methods listed above. This protocol was adapted using the manufacturer's instructions. Using a sample that had gone through DNase digestion or polyadenylation, the sample was adjusted to a volume of 100 µL with RNAse-free water. Next, 350 µL of Buffer RLT was added to the sample and mixed well. Then, 250 µL of ethanol was added and mixed well. The sample was transferred to a RNeasy Mini spin column with a collection tube and centrifuged for 15 seconds at 8000 x g. The

flow-through was discarded. Next, 500 µL of Buffer RPE was added to the RNeasy spin column and centrifuged for 15 seconds at 8000 x g. The flow-through was discarded. Next, 500 µL of Buffer RPE was added to the RNeasy spin column and centrifuged for 2 minutes at 8000 x g. The column was placed in a fresh tube and was centrifuged again for 1 minute at 8000 x g. The RNeasy spin column was placed in a new 1.5mL collection tube. Then, 30 µL of RNAse-free water was placed directly on the column, incubated at room temperature for 10 minutes, and centrifuged for 1 minute at 8000 x g. This last step was repeated. The sample(s) were stored ina\ a freezer at -80ºC for future use.

Polyadenylation

The Nanopore minION requires a poly(A) tail to sequence RNA. This protocol is adapted from New England BioLabs Poly(A) Tailing of RNA using *E. coli* Poly(A) Polymerase. This procedure was followed after completion of the RNA cleanup protocol, as the polyadenylation cannot be completed if EDTA is present. The following components were added to a microcentrifuge tube in the order listed in Table 2:

Component	Volume
RNA	1-10 µg in 15 µL nuclease-free water
10X *E. coli* Poly(A) Polymerase Reaction Buffer	2 µL (1X)
ATP (10mM)	2 µL
E. coli Poly(A) Polymerase	1 µL
Total	20 µL

Table 2: Materials Required for Polyadenylation. This table shows the materials required to add a poly(A) tail to our bacterial samples using New England BioLab's E. coli Poly(A) Polymerase.

The reaction was incubated at 37ºC for 30 minutes. The reaction was stopped by adding EDTA to a final concentration of 10mM. The RNA cleanup procedure (listed above) was followed again and a sample was run on a 0.9% agarose gel to validate.

KBase and Analysis

The goal of our transcriptomic analysis is to determine the expression levels of different genes based upon RNA reads. To this end we make use of a bioinformatics pipeline whereby we align the reads to a genome sequence and then assemble the reads; from the assembled reads we can determine the expression level of genes based on their FPKM value.

In the previous work by Harding, transcriptomic analysis was conducted using KBase, an online service developed by the U.S. Department of Energy as a way to simplify the creation of

and use of analysis pipelines. However, we ran into difficulty with KBase, specifically, we found that it took far too long to perform the analysis at the scale that we needed it to. To remedy this we crated a new pipeline that was run in a Conda environment. First, the sample reads were aligned using bowtie2, then the reads were assembled using stringtie. In order to test our new pipeline, analysis was carried out on studies previously analyzed by Harding. We found that our pipeline differed slightly in our results, however still maintained the same overall patterns as the pipeline previously used. The results of the pipeline as well as the studies used can be found in appendix F.

Command	
Make reference	usr$ bowtie2-build /path/to/reference.fa genome_name
Get SRA Reads	usr$ fastq-dump SRA_Accession
Align Reads	usr$ bowtie2 -x genome_name -U /path/to/reads.fq -S /path/to/eg1.sam
Sort Reads	usr$ samtools sort -o /path/to/eg1.sorted.bam /path/to/eg1.sam
Assemble Reads	usr$ stringtie -o output.gtf eg1.sorted.bam -G reference.gff3

Table 3: Commands and their equivalent code. The code is run in the python terminal.

Results

RNA extraction of *P. putida* was attempted through various methods, all with varying results. Each method was analyzed via agarose gel electrophoresis and PCR amplification. Each run through of the methods was performed in duplicate on previously created RNA preps, due to time limitations, although often each sample yielded comparable results. Through the end of the time spent in the lab, we had significant difficulty in obtaining a pure RNA sample without significantly hindering the sample's concentration. A small factor in this could be due to the fast degradation of RNA. The following results are of cells grown in SESOM media, labeled S1, although it is noted that the results of cells growing LB media had remarkably similar results.

Isolation of Pure RNA

In order to identify the transcriptome, we first need to isolate clean RNA that is free of genomic DNA (gDNA) contamination. After successfully lysing the cells, it was necessary to use the RNeasy Mini kit to purify the RNA. We then ran an agarose gel to look for signs of gDNA contamination. We noticed a high molecular weight band along with post-lysis cell debris as indicated in lane 3 of Figure 1. We therefore determined it would be necessary to perform a DNase elimination step.

To further eliminate gDNA contamination, we performed a DNase treatment on the samples. The DNase treatment had been adjusted and moved around in the process to assess contamination. We then ran a gel to determine if any gDNA contamination existed in the samples. We saw that lane 4 shown in Figure 1 contained two rRNA bands, 23S and 16S, while containing no visible gDNA contamination. While our sample appeared pure, it contained other debris from the extraction process.

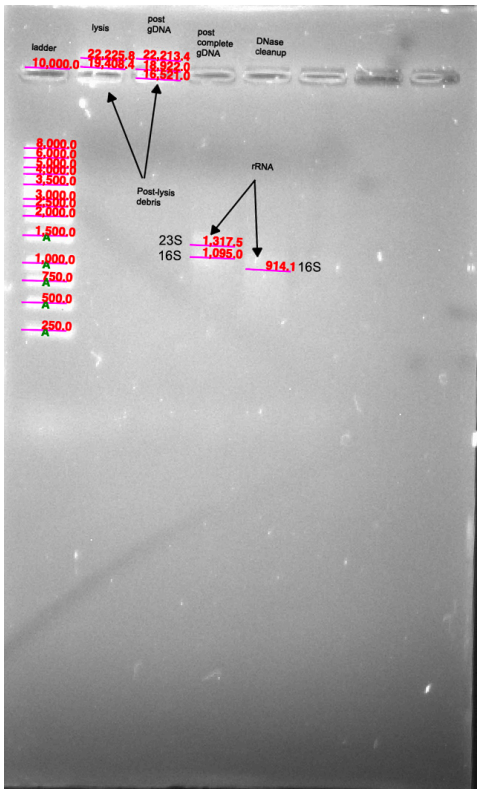


Figure 1: Agarose Gel Validation of gDNA Removal. This image shows 10μL of each sample mixture after validation by agarose gel electrophoresis. Some samples show evidence of gDNA contamination, while the further steps conclude rRNA bands are present without gDNA contamination.

In order to polyadenylate our samples, they cannot contain any traces of EDTA. To rid our samples of EDTA or any remaining impurities, we ran the samples through a clean-up column. We noticed a smaller 16S rRNA band appeared in lane 5 of Figure 1 after this step. To further verify that our samples did not contain any gDNA contamination, we performed a PCR reaction.

PCR Analysis

To further validate the removal of gDNA, a PCR analysis was performed. This was performed because evidence of gDNA contamination would be shown with a band since RNA cannot be amplified. We then ran our samples on a gel to detect any gDNA contamination. As a positive control in our gel, a sample of cDNA was loaded into lane 5 shown in Figure 2. Our gel showed that one of our samples (S2 found in lane 3 of Figure 2) contained gDNA contamination. Along with this, our gel showed evidence of primer dimers. One of our samples, S1, did not contain significant gDNA contamination and was therefore polyadenylated for sequencing purposes.

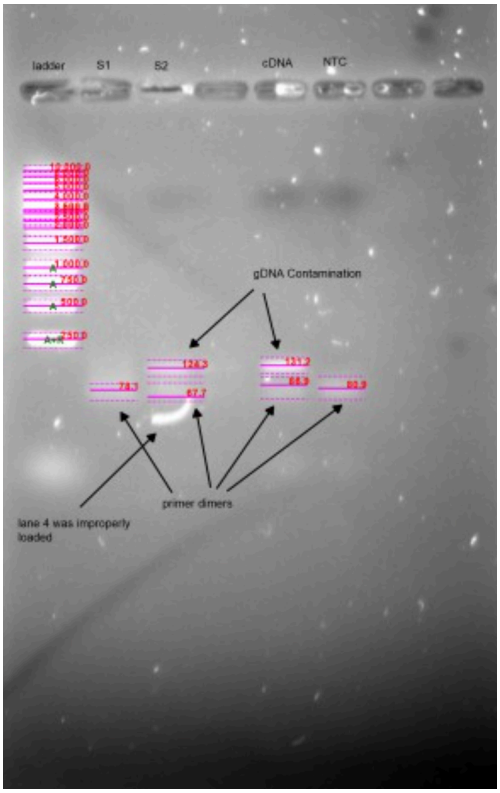


Figure 2: Agarose Gel Electrophoresis for the Validation of PCR Procedure. This image was taken using the GelDock and Image Lab Software. Lanes 2 and 3 contain our sample attempts at obtaining a sample of pure RNA. Evidence of primer dimers is noted in lanes 2, 3, 5, and 6. Our positive cDNA control was loaded into lane 5. A negative template control containing only primers and nuclease free water was loaded into lane 6. Lane 4 was improperly loaded and was disregarded in the analysis.

Polyadenylation

It is necessary to polyadenylate a purified RNA sample for Direct RNA sequencing as the Oxford Nanopore minION requires a poly(A) tail to sequence RNA. This is not found in P. putida, so a polyadenylation procedure had to be performed. This step was successfully performed on our first try for practice, as the sample we used did contain a high enough concentration to sequence. The validation of this is shown in Figure 3, where the bands in lane 2 are slightly larger compared to the sample without this step, ran in lane 3. It is noted here that it is possible to perform polyadenylation successfully in P. putida in order to further prepare the RNA sample for direct RNA sequencing.

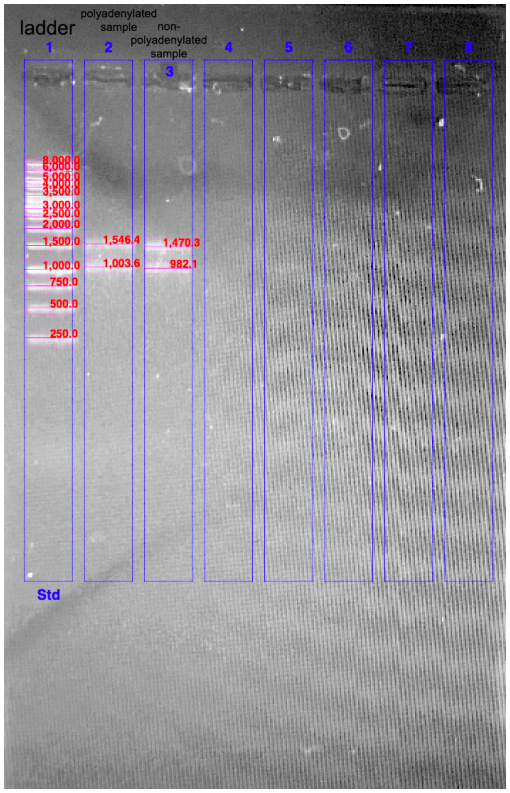


Figure 3: Agarose Gel Validation of Polyadenylation. Lane 2 holds the polydenylated sample, where lane 3 shows the purified sample pre-polyadenylated. Validation was proven using Image Lab software, where the bands show a 97 base pair difference where the poly-A tail was added.

Sample Concentration

Sample concentration is important to determination if a sample can be sequenced. It was determined that the sample, S1, had a concentration of 13.9 ng/µL, which is too low in order to sequence the RNA, where the input requirement is a minimum at least 500 ng RNA (poly-A+). Before sequencing, the total rRNA would need to be depleted, as well as putting the sample through a final sample cleanup column. This means that each sample did not meet the minimum requirements to proceed past the second cleanup phase.

Discussion

In order to analyze the transcriptome of P. putida KT2440, a pure sample of RNA was needed to be obtained for direct RNA-seq. Through various methods, we were able to obtain a pure sample of RNA. Along with this, we were able to successfully polyadenylate a bacterial sample for sequencing. It was discovered through the methods used to purify the RNA, the bacterial samples to have a low concentration.

RNA extracts of bacterial samples of P. putida KT2440 were successfully purified using the methods listed previously. We were able to validate these results through gel electrophoresis to ensure our samples did not contain gDNA contamination. This is because gDNA typically runs much slower through the gel pores than RNA.

PCR was a great tool to utilize in order to visualize if our samples contained any gDNA through amplification. This way, we could further validate our results through comparing our amplified samples with a positive and negative control and analyze if any bands were present.

While we were able to successfully polyadenylate one of our samples, it was merely for practice purposes. Our gel showed ~100 base pair difference between the polyadenylated sample vs the pre-polyadenylated sample. It is recommended that his step is performed after rRNA depletion to ensure the poly(A) tail is present before sequencing.

In our many attempts at validating our results, verification through gel electrophoresis proved to have some challenges. For example, primer dimers were evident in most gels, leading us to incorrectly believe the samples contained some type of contamination, including gDNA contamination. Due to this mistake, we would often discard our samples and we would start the RNA extraction and purification process over again because a cleanup column would further reduce the concentration. Similarly, through trial and error and research we concluded it is possible we accidentally UV bleached several of our gels while imaging, unsure on how long to expose the gels to obtain a clear image to analyze. Gel imaging challenges were also common when using PCR to validate our samples. Our team attempted using a 1% agarose gel as we had been for each validation, but the evidence was not clearly visible. Instead, a 0.9% agarose gel was poured and run to validate if polyadenylation was successful. The 0.9% gel proved to be the better choice for imaging and band separation.

The most challenging obstacle was the evidence of an extremely low concentration for each purified sample. After each purification step, the concentration of each sample would decrease significantly, especially using the cleanup columns, which decreased the concentration ten-fold. After each attempt to obtain a pure RNA sample, it was determined the concentration was too low to continue and each sample was then discarded. The methods we used to purify RNA have been typically used for cDNA library construction. The concentration of samples does not affect the creation of the library and was not considered as a possible challenge in the project until we were able to finally obtain a pure RNA sample. Modifications to our methods were attempted to account for this obstacle, but were not successful.

Through the transcriptomic analysis we were able to partially verify the results from the previous work done by Harding. However, we encountered some difficulty in determining whether or not either the previous work, or our new work was valid, as the papers that we got the RNA reads from do not publish the raw results of their analysis, only the conclusions that they

drew from it. Additionally, we recognized some of the samples were in fact paired ends, when they were incorrectly listed as single, so those results will need to be redone.

After several failed attempts at isolating pure RNA, we would like to suggest a few recommendations for the continuation of our project, assuming direct RNA-seq will still be utilized. First, we propose using a bacterial RNA retrieval kit, similar to the the PureLink™ Pro, which is manufactured by ThermoFisher. The PureLink™ Pro total RNA Purification Kit provides high-throughput isolation of high-quality total RNA. This kit is able to isolate high yields of total RNA with low genomic DNA contamination from bacteria. This is a one-hour procedure, compared to upwards of 15+ hours per sample, including validation, keeping in mind time dedicated in between classes and other schoolwork. The greatest benefit of this kit is that it allows for a high yield. There is a starting material maximum of up to 1 x 109 cells, with a yield of 33 µg. This is a promising kit, assuming after rRNA depletion the yield would be ≥ 500 ng RNA (poly-A+). The downside to this option is that each kit is awfully expensive, at about $1,700 per kit, although cheaper options with similar outputs may be available. This kit can be purchased and reviewed further at: https://www.thermofisher.com/order/catalog/product/12173011A?ICID=cvc-rna-bacteria-c5b1.

We also propose using Phenol:Chloroform to extract the RNA. This can replace the purification process we have previously mentioned, moving from culture production directly to PCR analysis. If contamination is not present, the samples can then move through the rRNA depletion process (assuming a high concentration) and the polyadenylation step for sequencing. This method also saves an ample amount of time. A suggested protocol for this step can be found in Appendix G. This protocol has been adapted from a published protocol by a team at the University of Colorado (Toni et al., 2018) and has not been performed using P. putida applying methods commonly used in Professor Farny's lab.

Our next recommendation applies to the NanoDrop UV-Vis Spectrophotometer that is found in Professor Young's lab. During our last attempt to isolate a pure RNA sample, we had noticed that the NanoDrop had never run its initial troubleshooting process back in 2017, or any other subsequent testing for accuracy. This machine had also not been updated since 2019, at least 3 years before the time of this project. We would like to note here that it is possible our concentration measurements could have been inaccurate from the start because of this. The NanoDrop has the capability of determining the concentration of both RNA and DNA in our samples, which could aid in the original protocol to detect gDNA contamination after each step and could reduce time in the validation process. Similarly, we used the concentration of samples to determine the amounts of varied materials used throughout the purification process, which could be affecting our yield.

Similar to the previous recommendation, we were able to identify protocols to estimate the amount of DNA and RNA in a given sample. The estimation of DNA is determined through a diphenylamine reaction, while RNA is estimated through an orcinol reaction (Jain et al., 2020). It is important to understand the true concentration of our samples, especially because the margin of error for the concentration is so small. The protocols for estimating DNA and RNA in samples can be found in Appendix H.

Finally, we recommend skipping the rRNA depletion step before sequencing. This would allow for a slightly higher concentration if the samples were to be on the threshold of 500 ng RNA (poly-A+). Since the 16S rRNA gene is the most highly conserved region amongst bacteria, it is easily identifiable, and the 23S,16S, and 5S rRNA sequences can be removed to leave the mRNA sequence.

References

Biolabs, New England. “NEBNext® RRNA Depletion Kit (Bacteria).” NEB. Accessed December 15, 2021. https://www.neb.com/products/e7850-nebnext-rrna-depletion-kit-bacteria#Quality,%20Safety%20&%20Legal_Specifications%20&%20Change%20Notifications.

Carvalhais, L. C., Dennis, P. G., Tyson, G. W., & Schenk, P. M. (2012). Application of metatranscriptomics to soil environments. Journal of microbiological methods, 91(2), 246-251.

D'Arrigo, Isotta, et al. "Analysis of Pseudomonas putida growth on Non- trivial Carbon Sources Using Transcriptomics and Genome-scale Modelling." Environmental Microbiology Reports, no. 2, Nov. 2018, pp. 87-97. doi:10:1111/1758-2229.12704

Frank, Sarah, et al. "Functional Genomics of the Initial Phase of Cold Adaptation of Pseudomonas putida KT2440." FEMS Microbiology Letters, no. 1, Mar. 2011, pp.47-54. doi:10.1111/j.1574-6968.2011.02237.x

Grünberger, F., Ferreira-Cerca, S., & Grohmann, D. (2022). Nanopore sequencing of RNA and cDNA molecules in Escherichia coli. RNA, 28(3), 400-417.

Harding, K. (2021). Transcriptomic Analysis of Pseudomonas putidafor Implementation of SoilMicrobial Biosensor (MQP).

Hewson, I., Poretsky, R. S., Dyhrman, S. T., Zielinski, B., White, A. E., Tripp, H. J., ... & Zehr, J. P. (2009). Microbial community gene expression within colonies of the diazotroph, Trichodesmium, from the Southwest Pacific Ocean. *The ISME journal*, *3*(11), 1286-1300.

Illumina (2022). Benefits of SBS Technology | Robust sequencing data quality. Illumina. https://www.illumina.com/science/technology/next-generation-sequencing/sequencing-technology/sbs-benefits.html

Jain, Aakanchha., Jain, Richa., Jain, Sourabh. (2020). *Basic Techniques in Biocehmistry, Microbiology and Molecular Biology: Principles and Techniques.* Springer Protocols. https://doi.org/10.1007/978-1-4939-9861-6

Jain, M., Olsen, H. E., Paten, B., & Akeson, M. (2016). The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome biology, 17(1), 1-11.

Kalyuzhnaya, M. G., Beck, D. A. C., Suciu, D., Pozhitkov, A., Lidstrom, M. E., & Chistoserdova, L. (2009, November 19). Functioning in situ: Gene expression in methylotenera mobilis in its native environment as assessed through transcriptomics. Nature News.

Martínez-García, E., & de Lorenzo, V. (2017). Molecular tools and emerging strategies for deep genetic/genomic refactoring of Pseudomonas. Current opinion in biotechnology, 47, 120-132.

Nanoporetech, (2021). Advantages of nanopore sequencing. Nanoporetech. https://nanoporetech.com/how-it-works/advantages-nanopore-sequencing

Nanoporetech, (2022). How nanopore sequencing works. Nanoporetech. https://nanoporetech.com/how-it-works

Oxford Nanopore Technologies (2021) RNA and gene expression analysis using direct RNA and cdna sequencing. Oxford Nanopore Technologies. (2021, October 26). Retrieved March 29, 2022, from https://nanoporetech.com/applications/rna-sequencing

Oxford Nanopore Technologies. (2022, March 4). How nanopore sequencing works. Oxford Nanopore Technologies. Retrieved March 29, 2022, from https://nanoporetech.com/how-it-works

Oxford, N.T. , (2022) RNA sequencing and gene expression analysis | Oxford Nanopore Technologies. Nanoporetech. https://nanoporetech.com/applications/rna-sequencing

Pitt, M. E., Nguyen, S. H., Duarte, T. P., Teng, H., Blaskovich, M. A., Cooper, M. A., & Coin, L. J. (2020). Evaluating the genome and resistome of extensively drug-resistant Klebsiella pneumoniae using native DNA and RNA Nanopore sequencing. Gigascience, 9(2), giaa002.

Poblete-Castro, I., Borrero-de Acuña, J. M., Nikel, P. I., Kohlstedt, M., & Wittmann, C. (2017). Host organism: Pseudomonas putida. Industrial Biotechnology: Microorganisms, 1, 299-326. Quainoo, Scott & Coolen, Jordy & van Hijum, Sacha & Huynen, Martijn & Melchers, Willem & Schaik, Willem & Wertheim, Heiman. (2017). Whole-Genome Sequencing of Bacterial Pathogens: the Future of Nosocomial Outbreak Analysis. Clinical Microbiology Reviews. 30. 1015-1063. 10.1128/CMR.00016-17.

Sarode, N., Parris, D. J., Ganesh, S., Seston, S. L., & Stewart, F. J. (2016). Generation and analysis of microbial metatranscriptomes. Manual of Environmental Microbiology, 2-4.

Shakya, M., Lo, C. C., & Chain, P. S. (2019). Advances and challenges in metatranscriptomic analysis. Frontiers in genetics, 10, 904.

Smith, A. M., Jain, M., Mulroney, L., Garalde, D. R., & Akeson, M. (2019). Reading canonical and modified nucleobases in 16S ribosomal RNA using nanopore native RNA sequencing. PloS one, 14(5), e0216709.

Thermo Fisher Scientific, (n.d.). *RNAlater faqs: Thermo Fisher Scientific - US*. RNAlater FAQs | Thermo Fisher Scientific - US. Retrieved April 29, 2022, from https://www.thermofisher.com/us/en/home/references/ambion-tech-support/rna-isolation/tech-notes/rnalater-faqs.html

Toni, L. S., Garcia, A. M., Jeffrey, D. A., Jiang, X., Stauffer, B. L., Miyamoto, S. D., & Sucharov, C. C. (2018). Optimization of phenol-chloroform RNA extraction. MethodsX, 5, 599-608.

Wangsanuwat, C., Heom, K. A., Liu, E., O'Malley, M. A., & Dey, S. S. (2020). Efficient and cost-effective bacterial mRNA sequencing from low input samples through ribosomal RNA depletion. BMC genomics, 21(1), 1-12.

Wilson, D. (2019, February 14). Direct RNA sequencing on nanopore arrays redefines the transcriptional complexity of a viral pathogen. Nature. https://www.nature.com/articles/s41467-019-08734-9

Zhao, L., Zhang, H., Kohnen, Markus V.., Prasad, Kasavajhala V. S. K.., Gu, L. & Reddy, Anireddy S. N.. (2019). Analysis of Transcriptome and Epitranscriptome in Plants Using

PacBio Iso-Seq and Nanopore-Based Direct RNA Sequencing. Front. Genet., 10. https://doi.org/10.3389/fgene.2019.00253

Appendices

Appendix A: Direct RNA- Sequencing vs Illumina Sequencing

Below are tables comparing the pros and cons of direct RNA-seq, which this project intended to use, and Illumina sequencing, one of the most widely used sequencing techniques used today.

Direct RNA Sequencing

Pros	Cons
Characterise and quantify full-length transcripts using long reads (Oxford, N.T. , 2022)	Truncated reads, especially missing nucleotides at the 5′ end of the transcripts (Zhao et al., 2019)
Get faster access to results with real-time analysis and bespoke tools (Oxford, N.T. , 2022)	Present bioinformatics tools for identification of RNA modification are rare (Zhao et al., 2019)
Explore epigenetic modifications through direct RNA sequencing (Oxford, N.T. , 2022)	Only polyadenylated RNAs will pass through the pore complex (Wilson, 2019)
Reduce bias with PCR-free protocols (Oxford, N.T. , 2022)	3’ bias due to sequence reads being generated 3′ − > 5′ (Wilson, 2019)
Individual polyadenylated RNAs are sequenced directly, without recoding and amplification biases (Wilson, 2019)	Each read maps with high specificity but comparatively low identity (80–90% (nanopore) vs. > 99.9% (Illumina)) to the reference genome/ transcriptome (Wilson, 2019)
Data is highly reproducible (Wilson, 2019)	Determining the breadth of protein variants is hindered by the presence of insertion/deletion and substitution-type errors within the raw nanopore reads (Wilson, 2019)
Produces ultra-long reads (Nanoporetech, 2022)	Sensitivity of biological nanopores to changes in experimental environment (Quainoo, 2017)
Real-time analysis (400 bases/second) (Nanoporetech, 2022)	Highest error rate of all platforms (Quainoo, 2017)
Nanopore devices are portable (Nanoporetech, 2021)	

Illumina Sequencing

Pros	Cons
Illumina reads are derived from highly fragmented poly(A)-selected RNAs (Wilson, 2019)	Challenges occur when applying short-read Illumina sequencing where distinct overlapping transcripts share the same poly(A) signal and RNA cleavage sites (Wilson, 2019)
Cost effective (Quainoo, 2017)	Produce short reads, often causing gaps (Nanoporetech, 2022)
Highest confirmed output (Quainoo, 2017)	Use of potentially biased DNA polymerases during bridge amplification (Quainoo, 2017)
Lowest error rates (Quainoo, 2017)	Long sequence runs (Quainoo, 2017)
Proven base calling accuracy (Illumina, 2022)	No real-time data access (Quainoo, 2017)
Best suited for outbreak protocols where high accuracy and reliability are prioritized (Illumina, 2022)	High instrument costs (Illumina, 2022)

Appendix B: Previous Relevant Studies Using Direct RNA-Sequencing on Bacterial Species

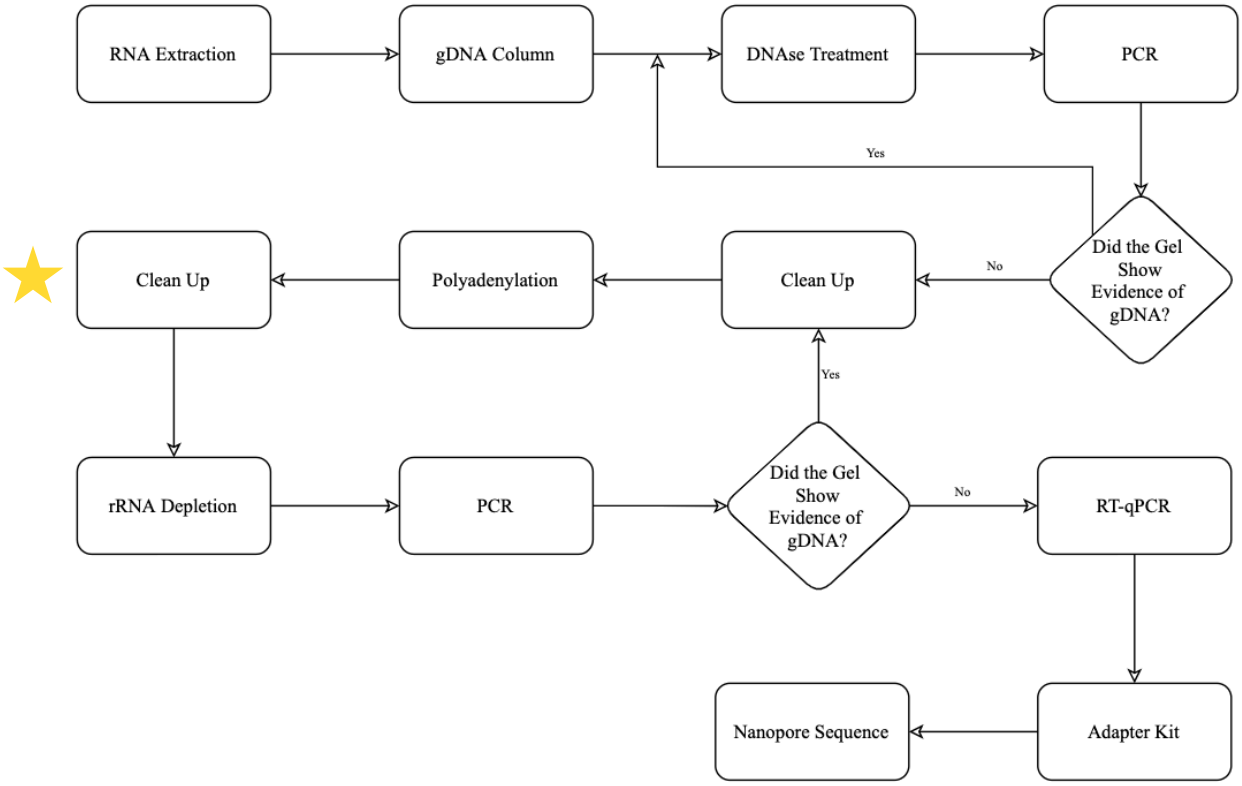
The table below shows the goal, bacterial strain, methods for RNA extraction and purification, and major findings of previous studies that successfully utilized direct RNA-seq on bacterial species

Goal	Strain	Methods	Major Findings
Implement the Direct RNA-seq technique for prokaryotic transcriptomes (Grünberger et al., 2022)	*E. coli* K-12 strain MG1655	Suggestions: • Ligate RNA to a double-stranded oligo(dT) splint adapter • Reverse transcription is highly recommended to resolve secondary structures in the RNA and to decrease the probability of pore blockage • rRNA depletion, However, the input quantity requirements currently still make it challenging to use rRNA-depleted RNAs (yield = 500 ng polyA+) Methods: • Purification of high-quality RNAs using silica-membrane columns with a cut-off size of about 200 nucleotides • Polyadenylation	This study presents an experimental and bioinformatic workflow for ONT RNA-seq in the bacterial model organism Escherichia coli, which can be applied to any microorganism.
Study the effectiveness of bacterial direct RNA sequencing and base-calling (Pitt et al., 2020)	XDR *K. pneumoniae* clinical strains	• RNA was extracted via the PureLinkTM RNA Mini Kit, using Homogenizer columns • To remove DNA contamination, the TURBO DNA-freeTM kit was implemented. A minor adjustment was an increased concentration of TURBO DNase (4 U) incubated at 37ºC for 30 min. • The RNeasy Mini Kit clean-up protocol was used to purify and concentrate RNA samples. • The rRNA was depleted via the MICROBExpressTM Bacterial mRNA Enrichment Kit. Minor protocol changes included adding ≥2 µg of DNA-depleted RNA, and the enriched mRNA was precipitated for 3h at −20ºC. • Poly(A) addition was performed using the Poly(A) Polymerase Tailing Kit • The input RNA concentration was ≥800 ng, and RNA samples were incubated at 37ºC for 1 hour.	Overall, MinION sequencing rapidly detected the XDR/ PDR K. pneumoniae resistome, and direct RNA sequencing provided accurate estimation of expression levels of these genes
Using direct sequencing of full-length 16S rRNA to overcome epigenetic modifications on rRNA that are erased by sequencing-by-synthesis methods (Smith et al., 2019)	*E. coli* strains BW25113 JW3718Δ and BW25113 JW2171Δ and *S. enterica* strain LT2	• *E. coli* strains K12 MG1655, RsmGΔ, RsuAΔ and *S. enterica* strain LT2 were grown in LB media at 37°C to an A_{600} = 0.8–1.0. • Cells were harvested by centrifugation and total RNA was extracted with Trizol • All total RNA samples were treated with DNase I (NEB) in the manufacturer's recommended buffer at 37°C for 15 minutes. • RNA was extracted by acid phenol/chloroform extraction and two rounds of chloroform extraction. • RNA was precipitated with sodium acetate (pH 5.2) and ethanol. • RNA was resuspended in nuclease-free water and stored at -80°C.	• Purified E. coli 16S rRNA was detected in total human RNA. • Conserved E. coli 16S rRNA 7-methylguanosine and pseudouridine modifications and a 7-methylguanosine modification that confers aminoglycoside resistance to some pathological E. coli strains.

Summarize the technical challenges for transcriptomics applied to soil environments and discuss approaches for gaining biologically meaningful insight into these datasets. (Carvalhais, 2012)	-	This review suggests using the following for mRNA enrichment: • Subtractive hybridization (MICROBExpress Bacterial mRNA Enrichment Kit, Ambion; Pang et al., 2004) • Exonuclease treatment, which preferentially degrades rRNA (mRNA-ONLY Prokaryotic mRNA Isolation kit, EPICENTRE Biotechnologies, Madison; USA) • Size separation by gel electrophoresis • Duplex specific nuclease (DSN) treatment	This review suggests challenges and methods for working with metatranscriptomics, specific to the soil environment.
Study the patterns of microbial gene utilization within colonies of Trichodesmium (Hewson et al., 2009)	*Trichode-smium*	• Cell pellets were removed and placed into RNase-free 2-ml cryovials, then subjected to the RNeasy Plant Mini kit • The resulting RNA was eluted in deionized water and subsequently treated to remove DNA using the RNase-free DNase kit • RNA was first subjected to terminator exonuclease treatment (which removes 50 -monophosphate-capped RNA) using the mRNA-ONLY protocol • rRNA was further reduced by subtractive hybridization using the MicrobExpress kit • The mRNA-enriched samples were amplified using in vitro transcription after mRNAs were polyadenylated, as part of the MessageAmp II—Bacteria aRNA kit • The aRNA (15 mg) was converted to double-stranded cDNA	The results provide insight into aggregate microbial communities in contrast to planktonic free-living assemblages that are the focus of other studies.

Appendix C: Flowchart of Methods Used in This Project

Details on how to perform each method can be found in the Methods section of this paper. Results on success can be found in the results section of this report. The results in further detail, including challenges presented and future recommendations, can be found in the Discussion section of this paper. The star on this flowchart indicates how far we were able to proceed with the methods before restarting with new RNA preps in hopes to gain improved results with modified results.



Appendix D: Protocol with Notes for Possible Modifications

Below is the entire protocol we followed to purify RNA samples. Along with this, there are several "†" indicating alternative protocols to consider, and "*" for specific notes about the section of the protocol

SESOM Creation

1. To a 500 mL Erlenmeyer flask, add 100 g of soil and 500 mL of 10X PBS buffer
2. Place in a 37ºC incubated shaker for 2-4 hours
3. Slowly pour the mixture into a French press to separate the liquid, decant slowly to ensure large solids stayed behind.
4. Place a piece of Cytiva Whatman circular qualitative filter paper Grade 1 (particle retention <11µm) under a piece of Grade 4 circular qualitative filter paper (particle retention <24µm) in a Büchner filtering flask funnel
    a. Saturate around the edges with PBS to ensure they are sealed
5. Leave the remaining liquid to be vacuum filtered overnight. Store in a 4ºC fridge for later use in culture growth.

Bacteria Sample Growth

LB Growth/Pellet

1. Add 5 mL of LB media to two (or however many you wish to make) 10 mL culture tubes
2. Inoculate the media with cultures of *P. putida* from the glycerol stock and incubate at 37ºC with shaking overnight
    a. Use a clean disposable inoculating loop each time
    b. Ensure the stock is not out if the -80ºC freezer for more than a few minutes
3. Back dilute the cultures by removing 0.5 mL of the culture and place the culture into a fresh 10 mL culture tube containing 5 mL of LB media.
4. Incubate at 37ºC for 90 minutes with shaking and placed in a 37ºC shaker for 90 minutes, or until the OD reaches ~0.45-0.55
5. Record the concentration and place 1 mL samples into 2 mL centrifuge tubes†
6. Centrifuge at 10,000 rpm for 10 minutes
7. Decant the supernatant††
8. Continue with purification††† or immediately store at -80ºC until further use

†From this step, it is possible to ignore the proceeding steps by using the protocol listed in Appendix G and validate gDNA removal through PCR, and proceed further with rRNA depletion and polyadenylation if contamination does not appear

††Alternatively, resuspended pellets in a small volume of PBS, then mixed with 5-10 volumes of RNAlater (available in the Farny lab). Once in RNAlater, samples can be stored for up to 1 day

at 37°C, 1 week at 25°C, 1 month or more at 4°C, and long-term at -20°C or -80° (Thermo Fisher Scientific, n.d.).

†††Alternatively, you can use a bacterial RNA retrieval kit and validate gDNA removal through PCR, and proceed further with polyadenylation if contamination does not appearCultured cells should be pelleted, resuspended in a small volume of PBS, then mixed with 5-10 volumes of RNAlater.

SESOM Growth/Pellet

1. Add 5 mL of SESOM media to two (or however many you wish to make) 10 mL culture tubes
2. Inoculate the media with cultures of *P. putida* from the glycerol stock and incubate at 37ºC with shaking overnight
    a. Use a clean disposable inoculating loop each time
    b. Ensure the stock is not out if the -80ºC freezer for more than a few minutes
3. When the OD reaches ~0.45-0.55, record the concentration and place 1 mL samples into 2 mL centrifuge tubes†
4. Centrifuge at 10,000 rpm for 10 minutes
5. Decant the supernatant
6. Continue with purification†† or immediately store pellets at -80ºC until further use

†From this step, it is possible to ignore the proceeding steps by using the protocol listed in Appendix G and validate gDNA removal through PCR, and proceed further with rRNA depletion and polyadenylation if contamination does not appear

††Alternatively, you can use a bacterial RNA retrieval kit and validate gDNA removal through PCR, and proceed further with polyadenylation if contamination does not appear

Cell Lysis

1. Clean lab benches and equipment with RNaseZap
2. Place 200 µL of lysozyme-TE buffer to centrifuge tubes containing a pellet.
3. Add 10µL of proteinase K to each tube. Pipette to resuspend the pellet
4. Mix by vortexing for 10 s
5. Incubate at room temperature for 10 min
    a. During incubation, vortex for 10 s every 2 min
6. Add 700 µL of Buffer RLT* and 7 µL of β-mercaptoethanol to each tube
7. Vortex for 10 s

* The adapted protocol states that Buffer RLT is only stable for 1 month after the addition of β-mercaptoethanol to the samples.

gDNA Removal and DNase Treatment

8. Place the lysate into a gDNA eliminator spin column
9. Centrifuge for 30 s at 8,000 x g
10. Treat 20 µL of extract at a time with 10X TURBO DNase Buffer to 1X concentration in the RNA sample*

11. Add 2 µL of TURBO DNase
12. Mix by inverting and incubate for 25 minutes at 37ºC in a dry heat bath†
13. Add a 15 mM concentration of EDTA and incubate in a dry heat bath at 75ºC for 10 min

*If the nucleic acid solution concentration is >200 µg/mL, dilute to 10 ng nucleic acid/50 µL.
†Alternatively, incubate for 30 min at 37ºC (Pitt et al., 2020)

gDNA Column

14. Determine the concentration of the RNA sample by using the Thermo Scientific™ NanoDrop™ One Microvolume UV-Vis Spectrophotometer (or similar)†
    a. Blanking with PBS
    b. Add 2 mL of sample and record concentration
15. Dilute to get 50 mL of a 100 ng/µL concentration
16. Add 500 µL of 98% ethanol and mix by pipetting
17. Add 700µL of sample to a RNeasy Mini spin column
18. Centrifuge for 15 s at 8000 x g
19. Discarded the flow-through
20. Add 700 µL of Buffer RW1 to the RNeasy spin column using the same collection tube.
21. Incubate at room temperature for 5 min
22. Centrifuge for 15 s at 8000 x g
23. Discard the collection tube containing the flow-through and place the column in a new collection tube
24. Add 500 µL of Buffer RPE with ethanol to the column
25. Centrifuge for 15 s at 8000 x g and discard the flow-through
26. Add 500 µL of Buffer RPE to the column
27. Centrifuge for 2 min at 8000 x g
28. Place the column in a fresh tube and centrifuge for 1 min at 8000 x g
29. Place the column was into a new 1.5 mL RNase-free collection tube
30. Add 30 µL of RNase-free water directly onto the column (do not touch column to avoid puncturing it)
31. Incubate at room temperature for 10 min
32. Centrifuge for 1 min at 8000 x g
33. To avoid lowering the concentration, place the flow-through on the column again
34. Incubate in room temperature for 10 min
35. Centrifuge for 1 min at 8000 x g

†For more accurate concentration determination, perform either a Diphenylamine reaction to estimate DNA or the Orcinol method for RNA estimation listed in Appendix H

PCR

36. Keep reaction components on ice for the duration of the process
37. Into a 0.2 mL PCR tube, add the following components:*

RNA Sample	+ Control	- Control
12.5µL of One*Taq* Quick-Load 2X Master Mix with Standard Buffer	12.5µL of One*Taq* Quick-Load 2X Master Mix with Standard Buffer	12.5µL of One*Taq* Quick-Load 2X Master Mix with Standard Buffer
0.5 µL of forward primer	0.5 µL of forward primer	0.5 µL of forward primer
0.5 µL of reverse primer	0.5 µL of reverse primer	0.5 µL of reverse primer
1µL of template DNA	1µl of template DNA (diluted to 1 µL cDNA in 4 µL of Nuclease-free water)	
10.5 µL of Nuclease-free water	10.5 µl of Nuclease-free water	11.5 µl of Nuclease-free water

*The primer set is for gene PP_2088 (sequence listed in Appendix E).

38. Mix the reaction mixture gently by pipetting.
39. Place samples in a thermocycler with the following conditions:

Step	Temperature	Time
Initial Denaturation	94ºC	30 seconds
30 Cycles	94ºC 67-68ºC 68ºC	20 seconds 20 seconds 15 seconds
Final Extension	68ºC	5 minutes
Hold	4ºC	∞

Gel Electrophoresis (for a 1% agarose gel)

40. Place a tared Erlenmeyer flask on an electronic balance
41. Weigh 0.5 g of agarose
42. Add 50 mL of 1XTAE
43. Tare the scale
44. Microwave the mixture for 1 min, swirl to mix
45. Microwave again for 1 min, swirl again
46. Microwave for 20 s more, swirl again
47. Place on the scale and add deionized water until the balance reads zero
48. Add of 2.5 µL of SYBR Safe and swirl to mix
49. Pour the agarose mixture into a gel casting tray with a well comb in place
50. Let sit at room temperature for 20-30 min
51. One the gel has set, place in the electrophoresis chamber filled with 1xTAE Buffer until the gel is covered.
52. Load the samples as follows:

Purified Samples	PCR Analysis	Polyadenylated Samples
2 µL of GeneRuler 1kb DNA Ladder	2 µL of GeneRuler 1kb DNA Ladder	2 µL of GeneRuler 1kb DNA Ladder
10 µL of each sample	5µL of each sample (including controls if using SYBR Safe)	10 µL of each sample (polyadenylated and non polyadenylated samples)

53. Run at 100 V for 30 minutes.
54. Image using the Bio-Rad GelDoc machine using UV light and analyze for gDNA contamination. To analyze band size, a free version of Image Lab software can be downloaded from the Bio-Rad website with a free account.

RNA Cleanup

55. After DNase digestion and polyadenylation, adjust the sample to a volume of 100 µL with RNAse-free water
56. Add 350 µL of Buffer RLT and mix well
57. Add 250 µL of ethanol and mix well
58. Transfer sample to a RNeasy Mini spin column with a collection tube
59. Centrifuge for 15 s at 8000 x g
60. Discard the flow-through
61. Add 500 µL of Buffer RPE
62. Centrifuge for 15 s at 8000 x g
63. Discard the flow-through
64. Add 500 µL of Buffer RPE
65. Centrifuge for 2 min at 8000 x g
66. Place the column in a fresh tube and centrifuge again for 1 minute at 8000 x g
67. Place the column in a new 1.5mL collection tube
68. Add 30 µL of RNAse-free water directly on the column (do not touch column to avoid puncturing it)
69. Incubate at room temperature for 10 min
70. Centrifuge for 1 min at 8000 x g
71. Add 30µL of flow-through directly on the column (do not touch column to avoid puncturing it)
72. Incubate at room temperature for 10 min
73. Centrifuge for 1 min at 8000 x g
74. Store sample(s) at -80ºC for future use, or proceed further

Polyadenylation

75. Ensure EDTA is not present (complete after cleanup procedure)
76. Add the following components to a 1.5 mL centrifuge tube in the order listed:

Component	Volume
RNA	1-10µg in 15µL nuclease-free water
10X *E. coli* Poly(A) Polymerase Reaction Buffer	2µL (1X)
ATP (10mM)	2µL
E. coli Poly(A) Polymerase	1µL
Total	20µL

77. Incubate at 37ºC for 30 min
78. Stop the reaction by adding EDTA to a final concentration of 10mM
79. To prepare samples for gel validation, add gel: 1 µL RNA poly(A)+ sample, 8 µL water, 1 µL SYBR Safe dye

These are the steps we were successfully able to complete for the duration of the project. Not included in this protocol is a protocol for RT-qPCR before sequencing. The protocol for the NEBNext rRNA depletion kit can be found here: https://www.neb.com/protocols/2019/09/18/protocol-for-rrna-depletion-using-nebnext-rrna-depletion-kit-bacteria-neb-e7850-neb-e7860. A complete protocol to sequence RNA using the Nanopore minION can be found here with a free account creation: https://community.nanoporetech.com/guides/minion/rna002/introduction. Similar protocol summaries can be found in Appendix B.

Appendix E: PCR Primer Set for *P. putida* KT2440 Gene PP_2088

F: 5’ - AGAGGCGTCCGAAGACAAGGCT - 3’

R: 5’ - ACGCAGCACCAGAATTTCCCGG - 3’

Appendix F: Transcriptomic Analysis

SRA Accession #	Study	*P. Putida* Strain	Read End Type	# of Reads	# of Reads (Labels)
SRP157937	D'Arrigo, Isotta, et al. "Analysis of *Pseudomonas putida* growth on Non-trivial Carbon Sources Using Transcriptomics and Genome-scale Modelling." *Environmental Microbiology Reports*, no. 2, Nov. 2018, pp. 87-97. doi:10:1111/1758-2229.12704	KT2440	Single	24	4 reads (Glucose, Citrate, Ferulic Acid, Serine)
SRP003636	Frank, Sarah, et al. "Functional Genomics of the Initial Phase of Cold Adaptation of *Pseudomonas putida* KT2440." *FEMS Microbiology Letters,* no. 1, Mar. 2011, pp.47-54. doi:10.1111/j.1574-6968.2011.02237.x	KT2440	Single	2	2 reads (10 degrees, 30 degrees)

Table of the studies used in the transcriptomic analysis pipeline. The reads can be found by searching the studies respective SRA accession numbers in the SRA read browser. The read labels denote the different conditions that the bacteria were grown under.

Gene names/ locus	Protein names	Glucose	Citrate	Ferulic Acid	Serine	Average log2 fold change
infC PP_2466	Translation initiation factor IF-3	7.4877	7.4100	7.441	7.6543	0.0151
galU PP_3821	UTP-glucose-1-phosphate uridylyltransferase (UDP-glucose prophosphorylase)	4.9585	5.2771	5.0654	4.9963	0.1544
PP_2088	RNA polymerase sigma factor SigX	4.6365	4.9205	4.7198	4.7043	0.1450
fagG PP_1914	2-oxoacyl-[acyl-carrier-protein] reductase	4.3294	4.4756	4.2615	4.4001	0.0496

Table of the expression levels of select genes taken from the assembled reads. The expression values under the columns Glucose, Citrate, Ferulic Acid and Serine are given in FPKM. The average log2 fold change was taken with respect to Glucose.

Gene names/ locus	Protein names	10 Degrees	30 Degrees	Log2 fold change
infC PP_2466	Translation initiation factor IF-3	11.9225	12.1253	0.2028
galU PP_3821	UTP-glucose-1-phosphate uridylyltransferase (UDP- glucose prophosphorylase)	7.4209	7.9157	0.4947
PP_2088	RNA polymerase sigma factor SigX	7.4609	8.0259	0.5650
fagG PP_1914	2-oxoacyl-[acyl-carrier-protein] reductase	7.9098	7.9593	0.0495

Table of the expression levels of select genes taken from the assembled reads. The expression values under the columns 10 Degrees and 30 Degrees are given in FPKM. The log2 fold change was taken with respect to 10 Degrees.

Appendix G: Protocol for Phenol: Chloroform RNA Extraction

This protocol has been gathered from the article Optimization of phenol-chloroform RNA extraction (Toni et al., 2018) and adapted for potential future use with this project.

Required Reagents and Equipment

RNase Zap
QIAzol (or other proprietary phenol-based reagent)
Chloroform
Isopropanol
Ethanol (EtOH)
Nuclease-free water
2 mL centrifuge tubes
1 mL pipette, 0.2 mL pipette and tips
Refrigerated centrifuge or centrifuge chilled to 4ºC
65ºC heat block

Procedure

Preparation

1. Prepare a clean area, using RNase Zap spray or other chemical decontaminant.

Homogenization and RNA extraction

2. Add 1 mL of cell culture to a 2 mL centrifuge tube and add 500 µL of QIAzol to each tube.*

> Note: 500 µL of QIAzol is sufficient for 100,000–800,000 cells. If more QIAzol is necessary, volumes of chloroform and isopropanol in subsequent steps should be adjusted accordingly.

3. Let sit at room temperature (RT) for 3 min.
4. Add 100 µL RNase-free chloroform to each tube containing 500 µL QIAzol.
5. Shake vigorously by hand for 15 s.
6. Let sit at RT for 3 min.
7. Centrifuge at 4ºC ≤ 12,000 x g for 15 min.
8. Add 100 µL chloroform to a new 1.5 mL tube.
9. Transfer RNA-containing upper aqueous phase (clear supernatant) into chloroform.
10. Repeat chloroform extraction 1 time:
    a. Shake vigorously by hand for 15 s.
    b. Let set at RT for 3 min.
    c. Centrifuge at 4ºC ≤ 12,000 x g for 15 min.

RNA Precipitation

11. Add 250 µL RNase-free isopropanol to a new 1.5 mL tube.

12. Transfer RNA-containing upper aqueous phase (clear supernatant) into isopropanol.
13. Invert by hand 10–20 times to mix.
14. Let sit at RT for 10 min.
15. Centrifuge at 4ºC ≤ 12,000 xg for 10 min to precipitate RNA.

    Note: There should be a visible small white pellet following precipitation, however if RNA concentration is very low, a pellet may not be visible.

16. Remove supernatant and discard.

RNA Wash

17. Add 1 mL of 75% EtOH in nuclease-free water to pellet.
18. Centrifuge at 4ºC ≤ 7500 x g for 5 min.
19. Remove supernatant and discard.
20. Repeat EtOH wash 2 times:
    a. Add 1 mL of 75% EtOH in nuclease-free water to pellet.
    b. Centrifuge at 4ºC ≤ 7500 x g for 5 min.
    c. Remove supernatant and discard.
21. Vortex samples at RT.
22. Carefully remove remaining supernatant with pipette without disturbing the RNA pellet
23. Leave tubes open at RT for 3–5 min to evaporate EtOH.

    Note: Alternatively, excess EtOH can be removed by carefully tipping tubes upside down onto a Kimwipe.

RNA Solubilization

24. Add 20 mL of nuclease-free water to the RNA pellet.
    Note: Volume of water to be added can be optimized for desired RNA concentration. Heat tubes at 65ºC for 2–5 min to solubilize RNA.
25. Vortex tubes 5–10 s, vortex, and place solubilized RNA on ice immediately.
26. Quantify RNA concentration and purity.
27. Use RNA for downstream applications or freeze at -80 C immediately.

*Indicates methodological deviations from the conventional phenol-based RNA extraction protocol.

Appendix H: Estimation of DNA and RNA in Samples

The following protocols have been taken from Basic Techniques in Biochemistry, Microbiology and Molecular Biology: Principles and Techniques by Aakanchha Jain, Richa Jain, and Sourabh Jain.

Estimation of DNA by Diphenylamine Reaction

Reagents

1. DNA standard (100 mg/mL).
2. Standard citrate saline buffer (0.15 M sodium chloride +0.015 M sodium citrate).
3. To make diphenylamine reagent (DPAR): dissolve 1.5 g diphenylamine in 100 mL of acetic acid, add 2.75 mL of concentrated sulfuric acid. Store reagent at 2 degrees C.

Procedure

1. Prepare 100 mg/mL DNA standard solution and unknown isolated DNA sample in citrate buffer.
2. Prepare 3 mL aliquots of DNA standard (50-500 µg/mL)
3. Add 6 mL of DPAR to each aliquot. Mix and heat tubes in boiling water bath for 10 min then cool.
4. Using citrate buffer as blank, take the absorbance at 595 nm by UV spectrophotometer.
5. Create a standard curve from the known samples and predict the absorbance of the unknown sample.

Estimation of RNA Using Orcinol Method

Materials and Reagents

1. Standard RNA solution.
2. Orcinol solution (6% w/v in 95% ethanol).
3. Bial's reagent/orcinol acid reagent (10% w/v ferric chloride hydrated 1 mL added to concentrated hydrochloric acid 200 mL).
4. Tris-acetate buffer (1 mM EDTA added to 10 mM tris-acetate pH 7.2) or any buffer of pH 7.2.

Procedure

1. Dissolve 100 mg RNA pure standard in ice-cold tris-acetate buffer.
2. Prepare 3 mL aliquots of RNA standard (50-500 µg/mL) using RNAse-free water.
3. Add 6 mL of orcinol acid reagent to each aliquot. Mix and add 0.4 mL of orcinol solution. Mix again.
4. Heat the tubes in boiling water bath for 20 min then cool.
5. Create a blank by combining the previously used ingredients in RNAse-free water,
6. Using the previously created blank, take the absorbance at 665 nm by UV spectrophotometer.
7. Create a standard curve from the known samples and predict the absorbance of the unknown sample.