diagenode

Innovating Epigenetics Solutions

D-Plex Total RNA-seq Kit

Total RNA library preparation kit for Illumina® sequencing

Cat. No. C05030031 (24 rxns)

USER GUIDE

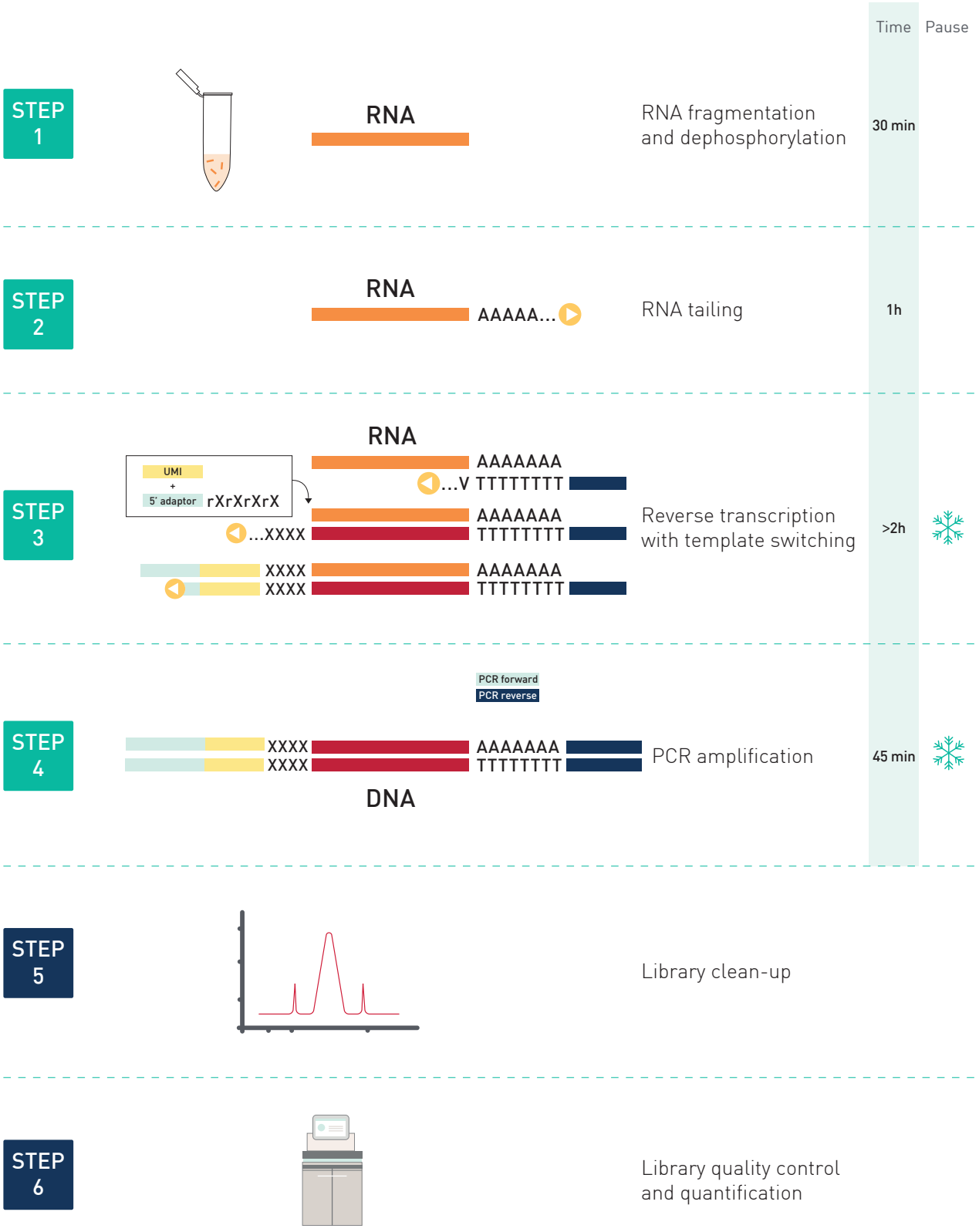
V1 05_2021

Please read this manual carefully before starting your experiment

Indexes are not included in this kit and are available separately. Please check "Required materials not provided" section.

Summary

Kit Method Overview



Introduction

The Diagenode D-Plex Total RNA-seq Library Preparation kit is a tool designed for the study of the **whole coding and non-coding transcriptome**. The present kit incorporates the unique **D-Plex technology** to generate directional RNA libraries for Illumina sequencing directly from total RNAs, messenger RNAs (mRNAs) that have already been enriched by poly(A) capture, or RNAs that have already been depleted of ribosomal RNAs (rRNAs).

The D-Plex technology utilizes the innovative **capture and amplification by tailing and switching**, a ligation-free method for library preparation and offers key advantages such as:

- **Ultra-low input** capability of the library preparation
- Ease of use in a **one day, one tube protocol**
- **High quality libraries** with **higher library complexity** than most of other available library preparation kits for RNA-sequencing

The library preparation protocol delivers a **high-sensitivity detection** method that captures the widest possible **diversity of RNAs** even from the most **challenging samples**. The kit enables the production of **high-quality libraries** within a wide range of RNA quantities going **from 50 pg up to 50 ng** of either purified total RNAs, poly(A)-isolated mRNAs or rRNA-depleted RNAs. The protocol is highly suitable for both intact and highly degraded RNA samples, including that derived from **FFPE preparations**.

The core of the technology relies on **ligation-free reactions** to attach the Illumina adaptors to both ends of the library construct. The entire library construction protocol, starting with total RNAs, can be completed in less than 5 hours with minimal handling steps and hands-on time. The workflow includes template fragmentation and end repair, polynucleotide tailing, reverse transcription with template-switching, and PCR amplification. The directionality of the template-switching reaction preserves the **strand orientation** of the RNA template, which thus provides indication on the DNA strand from which the RNA was transcribed. Retaining strandness information allows for identification

of antisense transcripts, determination of the transcribed strand of non-coding RNAs, and measurement of expression levels of coding or non-coding overlapping transcripts.

For optimal workflow flexibility, D-Plex Unique Dual Indexes Modules (C05030021 and C05030022) are available separately from the library preparation kit, providing PCR primers for library multiplexing up to 48. The use of **unique dual indexes (UDI)** is highly recommended to mitigate errors introduced by read misassignment, including index hopping frequently observed with patterned flow cells such as Illumina's NovaSeq 6000 system.

An important addition to the D-Plex set of features is the use of **unique molecular identifiers (UMI)** to each transcript incorporated in the library. Given this new addition, it is now possible to exclude PCR duplicates from a set of reads, thus improving the transcript expression quantification.

Kit Materials

Table 1. Components of the D-Plex Total RNA-seq kit (C05030031)

Component	Cap color	Qty (24 rxns)	Storage
Fragmentation Buffer (FB)	Yellow	24 µl	-20°C/-4°F
Dephosphorylation Reagent (DR)	Yellow	7.2 µl	-20°C/-4°F
Dephosphorylation Buffer (DB)	Yellow	48 µl	-20°C/-4°F
Tailing Reagent (TR)	Red	12 µl	-20°C/-4°F
Tailing Buffer (TB)	Red	24 µl	-20°C/-4°F
Reverse Transcription Reagent (RTR)	Purple	24 µl	-20°C/-4°F
Reverse Transcription Buffer (RTB)	Purple	120 µl	-20°C/-4°F
RT Primer H **UDI** (RTPH_UDI)	Purple	24 µl	-20°C/-4°F
RT Primer M **UDI** (RTPM_UDI)	Purple	24 µl	-20°C/-4°F
Template Switching Oligo **UDI** (TSO_UDI)	Purple	48 µl	-20°C/-4°F
PCR Master Mix (PCRMM)	Green	1200 µl	-20°C/-4°F
Nuclease-free Water	Clear	2000 µl	Room T°
Positive Control HBR (CTL+)	Black	12 µl (1 ng/µl)	-20°C/-4°F Long term storage: -80°C/-112°F

Storage

The components should be stored at temperatures indicated in Table 1.

Required Materials Not Provided

Indexes

Specific D-Plex indexes were designed and validated to fit this technology and are **not included** in this kit, providing you with total flexibility. They can be bought separately according to your needs. Please choose the format that suits you best among the compatible references to:

Table 2. D-Plex Unique Dual Indexes Modules and their corresponding references

Unique Dual Indexes Modules	Format	Reference
D-Plex Unique Dual Indexes Module – Set A	24 rxns	C05030021
D-Plex Unique Dual Indexes Module – Set B	24 rxns	C05030022

General equipment and reagents

- Gloves to wear at all steps
- Single channel pipettes and corresponding RNase-free filter tips: 10 µl, 20 µl, 200 µl, 1,000 µl
- Crushed ice
- RNase AWAY™ decontamination reagent (ThermoFisher Scientific, 10328011)
- Nuclease-free tubes: 0.2 ml, 1.5 ml
- Table top centrifuge with strip rotor
- Vortex agitator
- Tube holder for 0.2 ml, 1.5 ml tubes
- DiaMag 0.2 ml tube magnetic rack (Diagenode, B04000001)
- DiaMag 1.5 ml tube magnetic rack (Diagenode, B04000003)
- Thermal cycler

RNA enrichment before library preparation

- For enrichment of mRNA transcripts, we recommend using the **D-Plex mRNA-seq Kit** (C05030033) that contains a complete solution for gene expression analysis, or alternatively order the **D-Plex mRNA Capture Module** (C05030032).
- For depletion of rRNA transcripts, we recommend using a RNase H-based depletion method (e.g. NEBNext® rRNA Depletion Kit (NEB, E7400 or E7405))
- Absolute ethanol (VWR, 20821.310)
- Agilent 2100 Bioanalyzer® (Agilent) and Agilent RNA Pico 6000 (Agilent, 5067-1513)

Post-PCR libraries clean-up

- Agencourt® AMPure® XP Beads (Beckman Coulter, A63881)
- Absolute ethanol (VWR, 20821.310)
- DNase-free, RNase-free ultrapure water

DNA library size and yield estimation

- Agilent 2100 Bioanalyzer® Agilent and High Sensitivity DNA Kit (Agilent, 5067-4626)
- Qubit® Fluorometer (Thermo Fisher Scientific) and Qubit® dsDNA HS Assay Kit (Thermo Fisher Scientific, Q32851)

Remarks Before Starting

General recommendations

- Read the complete manual before first time use.
- Decontaminate the working area as well as all the tools used to perform the library preparation with RNase AWAY™ reagent.
- Wear gloves at all steps to protect the RNA sample from degradation by contaminants and nucleases.
- All containers and storage areas must be free of contaminants and nucleases.
- Add enzymes to reaction solutions last and thoroughly incorporate them by pipetting up and down the solution several times, and never vortex reagents mix.
- When using the protocol for the first time for a definite type of sample, we strongly recommend using the positive HBR control included in the kit.
- RNA samples with chemical modifications (e.g. from FFPE tissue) usually generate lower library yields and sequencing quality, so we recommend using more than 10 ng as starting amount before any RNA enrichment procedures.
- It is not recommended to make libraries using samples with a RIN < 2 or a DV200 < 20.
- Unless otherwise stated, the lid of the thermal cycler used to carry the reactions should be set at 105°C.

Designing your D-Plex RNA-seq experiment

The D-Plex Total RNA-seq protocol incorporates various RNA biotypes (including coding RNA, pseudogenes, tRNA, rRNA, long non-coding RNA, and small non-coding RNA) in the final library for sequencing.

If you work with already purified or enriched RNAs and you do not intend to perform an enrichment for mRNA transcripts or depletion of rRNA transcripts, please use the D-Plex Total RNA-seq kit (C05030031) and proceed with the library preparation protocol.

If you want to focus your whole transcriptome analysis to functionally relevant coding and non-coding transcripts, we recommend removing highly abundant rRNA species before proceeding with the D-Plex Total RNA-seq library preparation protocol. In this case, **at the end of the rRNA depletion protocol**, please pay attention that **the final elution of the RNA sample should be done in 7 µL of nuclease-free water**. We recommend using an RNase H-based depletion method (e.g. NEBNext® rRNA Depletion Kit (NEB, E7400 or E7405)). It is worth noting that by studying the complete transcriptome a considerable proportion of intronic reads should be expected since transcripts both before and after splicing are included in the library.

If you are specifically interested in gene expression analysis, we recommend using the **D-Plex mRNA-seq kit** (C05030033), or alternatively order the **D-Plex mRNA Capture module** (C05030032).

If you are specifically interested in small non-coding RNA discovery and profiling, we recommend using the **D-Plex Small RNA-seq kit** (C05030001).

Therefore, the choice of the D-Plex RNA-seq kit should be made based **on a clear understanding of the scientific question** in a given environment before proceeding to RNA-seq library preparation as the choice of methodology will strongly impact the end results.

Template

The D-Plex Total RNA-seq kit has been developed for efficient library preparation from high-quality, partially degraded, or low-quality total RNAs (RIN ≥ 2 and or a DV200 ≥ 20), including total RNAs extracted from FFPE tissues.

The D-Plex Total RNA-seq kit has also been validated on isolated poly(A) mRNAs and rRNA-depleted RNAs. After such enrichment, if the quantity of recovered RNAs is not limited, we recommend to assess the RNA concentration, integrity and quality using the Agilent RNA 6000 Pico kit (Agilent, 5067-1513) before library preparation. This is only possible if the starting amount of total RNAs is high enough to recover sufficiently enriched RNAs which can then be detected by the BioAnalyzer® system.

If quantification is not possible because of limited amount of material, we recommend using a control sample at high concentration in parallel to assess a percentage of recovery. This recovery rate (usually 0.1-1% for poly(A) mRNAs and 1-5% for rRNA-depleted RNAs) can be calculated and taken into account to estimate the quantity of RNA used for library preparation.

Table 3. Estimation of the quantity of recovered RNAs based on the enrichment methods

Starting amount	Expected amount	
Total RNAs	**poly(A) mRNA capture**	**rRNA depletion**
1 µg	1 – 10 ng	10 – 50 ng
100 ng	100 pg – 1 ng	1 – 5 ng
50 ng	50 pg – 500 pg	0.5 – 2.5 ng
10 ng	10 pg – 100 pg	100 pg – 500 pg

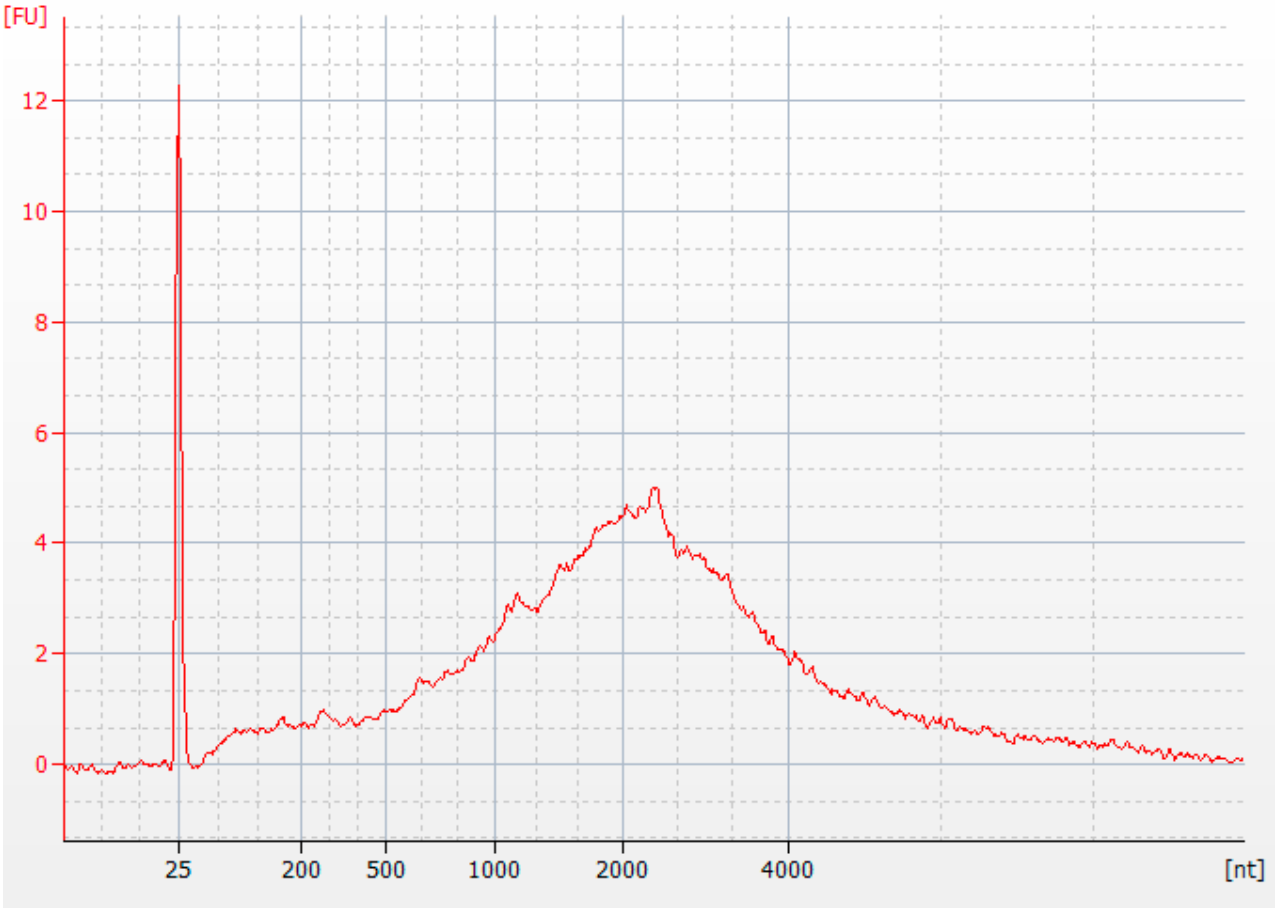


***Figure 2**: RNA electrophoregram of poly(A) selected mRNAs after the mRNA capture from 1 µg total Human Brain RNA (HBR) using the D-Plex mRNA Capture module (Diagenode, C05030032).*

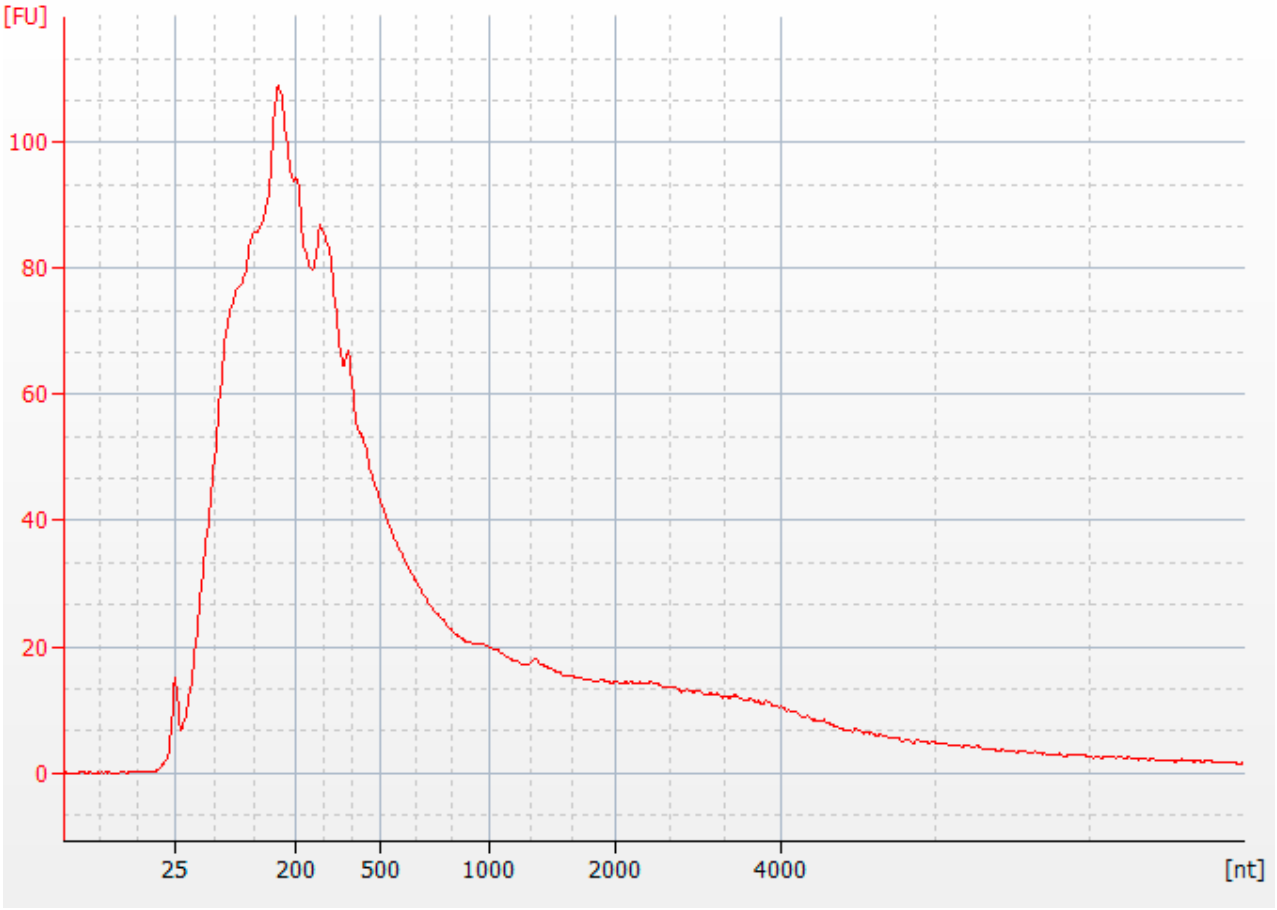


***Figure 3**: RNA electrophoregram of rRNA- depleted RNAs after the rRNA depletion from 1 µg total Human Brain RNA (HBR) using NEBNext® rRNA Depletion Kit v2 (NEB, E7400).*

Starting material

The D-Plex Total RNA-seq protocol has been validated for starting amounts ranging **from 50 pg up to 50 ng** of either purified total RNAs, isolated poly(A) mRNAs or rRNA-depleted RNAs. However, as RNA samples with chemical modifications (e.g. extracted from FFPE tissues) usually generate lower library yields and sequencing quality, we highly recommend starting the library preparation with 500 pg of RNAs for degraded samples. Increasing input quantity will improve library complexity.

A **starting volume of 7 µl** is used in the library preparation process. RNA templates should be dissolved in nuclease-free water in this volume before starting the protocol.

The RNA sample should be free of salts (Mg2+, guanidinium salts) or organic compounds (phenol, ethanol).

Multi-sample protocol

The protocol describes the library preparation process for one single sample. If the interest is to prepare more than one sample in parallel, please scale up accordingly.

Multiplexing advices

The PCR primers can be used for library multiplexing up to 48. In case of a multiplexing scenario, it is recommended to follow Illumina's library pooling guidelines that are explained in the D-Plex Unique Dual Indexes manual.

Positive control: total human brain RNA

A positive control is supplied in the kit for 12 reactions at a ready-to-use concentration of 1 ng/µl. This control RNA is meant to monitor the library preparation efficiency. Typically, 1 ng of control RNA prepared with the D-Plex protocol and amplified during 14 PCR cycles will yield a minimum of 5 ng/µl of dsDNA after library purification in 20 µl.

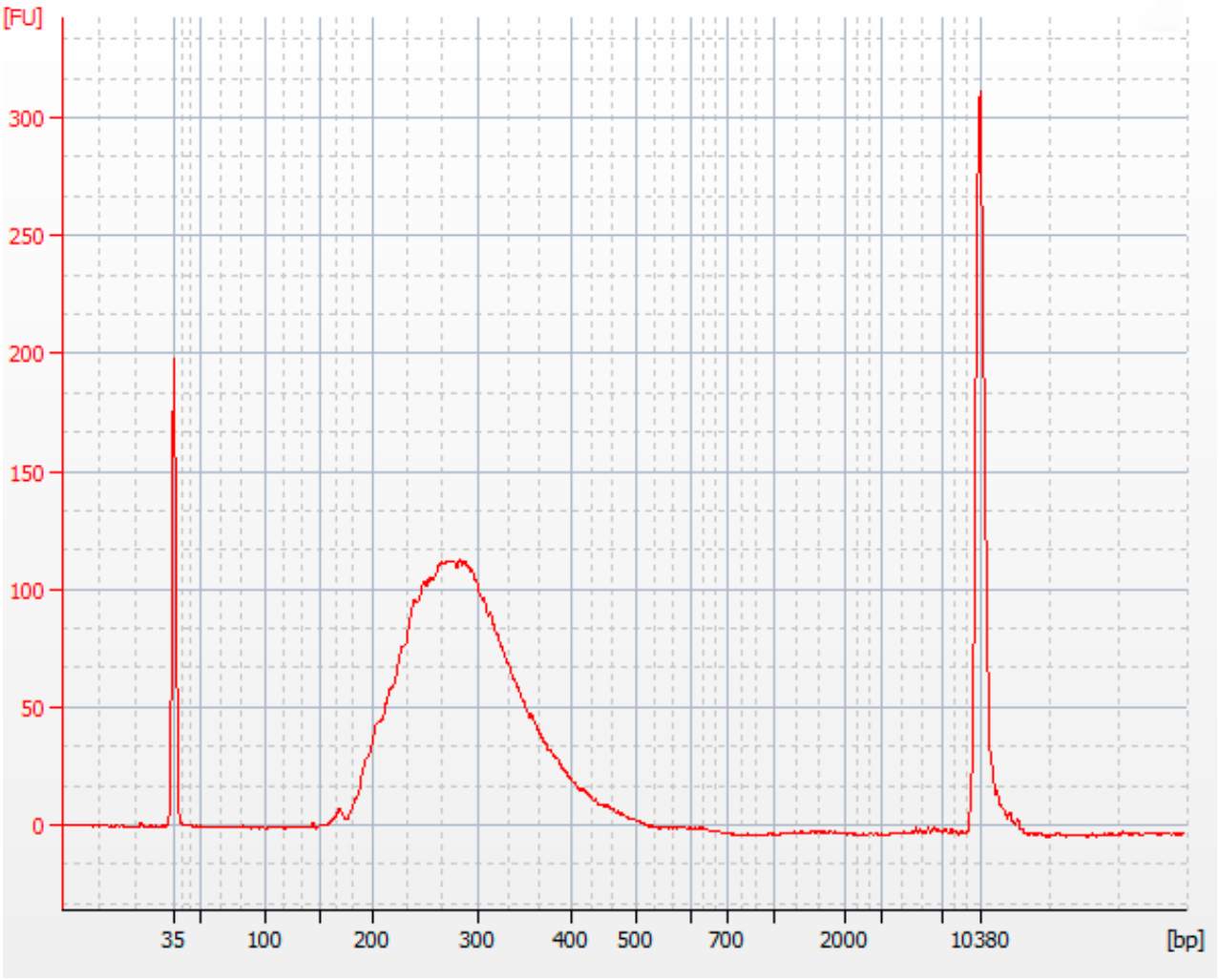
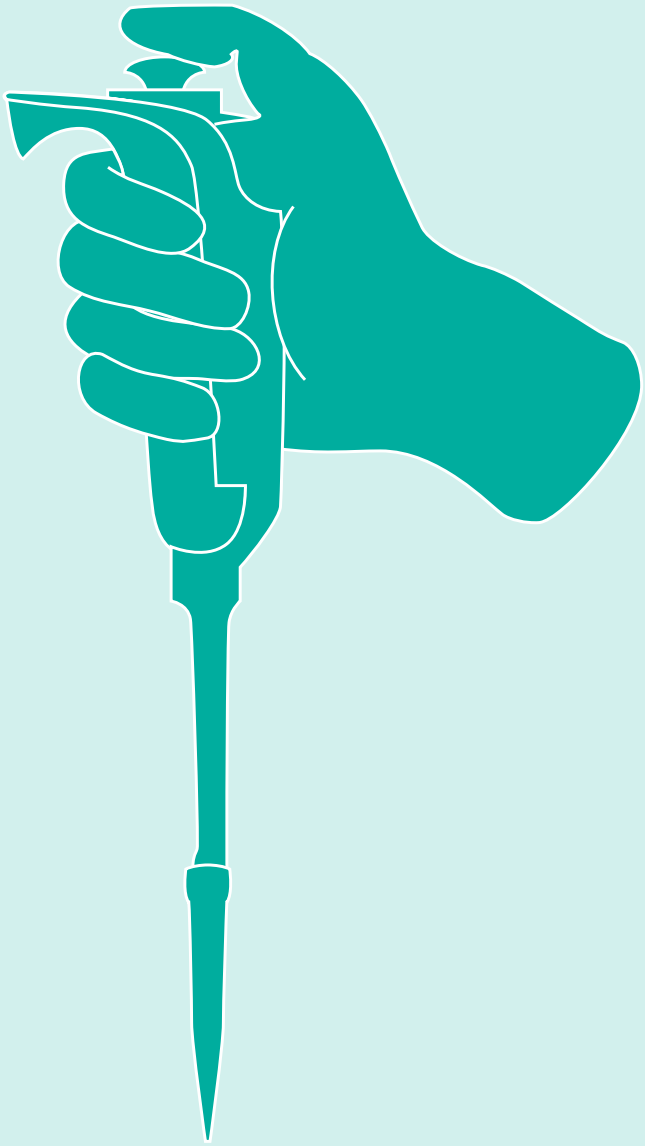


Figure 4: DNA electrophoregram of a D-Plex Total RNA-seq library prepared from 1 ng total Human Brain RNA (HBR). 1 ng of the library was loaded on the Bioanalyzer.

SHORT PROTOCOL



FOR EXPERIENCED USERS

Short Protocol For Experienced Users

***Starting material**: 50 pg – 50 ng of either purified total RNAs, isolated poly(A) mRNAs or rRNA-depleted RNAs in a 7 µl total volume of nuclease-free water.*

1. Add **1 µl of Fragmentation Buffer (FB)** to the **7 µl of the RNA sample**, incubate for 7 minutes at 94°C, and then cool down on ice.

   ***NOTE**: For highly degraded RNA samples (e.g., from FFPE tissues), DO NOT incubate at 94°C. Instead, incubate on ice for 7 minutes.*

2. Prepare the **Dephosphorylation Master Mix (DMM)** by mixing **2 µl of Dephosphorylation Buffer (DB)** and **0.3 µl of Dephosphorylation Reagent (DR)** in a RNase-free tube.

3. Add **2 µL of the DMM** in the sample, incubate for 15 minutes at 37°C, and then cool down on ice.

4. Prepare the **Tailing Master Mix (TMM)** by mixing **1 µl of Tailing Buffer (TB)** and **0.5 µl of Tailing Reagent (TR)** in a RNase-free tube.

5. Add **1.5 µl of TMM** to the sample, incubate for 40 minutes at 37°C + 20 minutes at 65°C, and then cool down on ice.

6. Add **1 µl of Reverse Transcription Primer UDI (RTP_UDI)** to the sample on ice:

   - For 50 ng - 500 pg RNA, use **RTPH_UDI**
   - For 500 pg - 50 pg RNA, use **RTPM_UDI**

7. Incubate for 10 minutes at 70°C, then slowly decrease to 25°C by ramping down at 0.5°C/second, and incubate for 2 minutes at 25°C.

8. Take the **Template Switch Oligo UDI (TSO_UDI)** tube out of the freezer and let it thaw on ice.

9. Prepare the **Reverse Transcription Master Mix (RTMM)** by mixing **5 µl of Reverse Transcription Buffer (RTB)** and **1 µl of Reverse Transcription Reagent (RTR)** in a RNase-free tube. Place the RTMM on ice until addition to the sample.

10. Add **6 µl of RTMM** to the sample and incubate for 15 minutes at 25°C.

11. Add **2 µl of TSO_UDI** to the sample, and incubate for 120 minutes at 42°C + 10 minutes at 70°C.

12. Add **20 µl of D-Plex Primer UDI** and **50 µl of PCR Master Mix (PCRMM)** to the sample.

13. Proceed to PCR reaction and amplify the library according to the following program:

Temperature	Time & Cycles	
98°C	30 seconds	
98°C	15 seconds	n cycles*
72°C	60 seconds	
72°C	10 minutes	
Hold at 4°C or freeze until further processing		

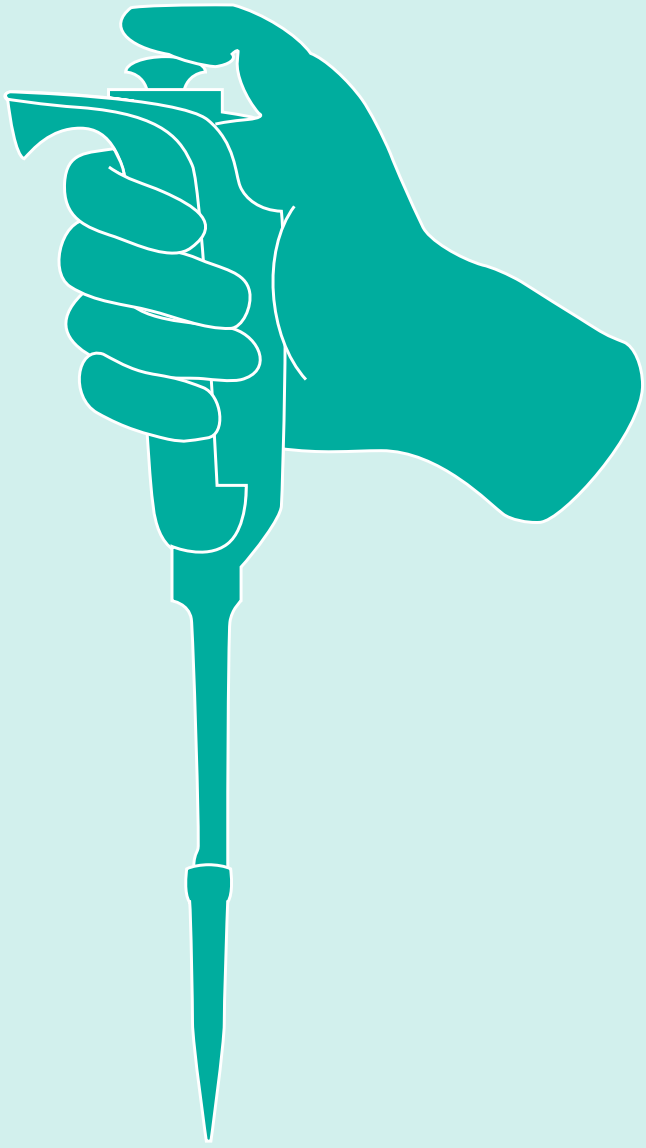
**Perform* n *cycles depending on the initial RNA input.See detailed protocol for the PCR cycle table.*

14. Clean-up the PCR product with **72 µL of AMPure® XP beads (0.8X)** and elute the purified library in **50 µl of nuclease-free water**.

15. Perform a second clean-up with **40 µL of AMPure® XP beads (0.8X)** and elute the final library in **10-20 µl nuclease-free water** (see detailed protocol for more information).

16. Perform library quantification using QuBit® dsDNA HS Assay Kit and quality check using Bioanalyzer® DNA High Sensitivity Assay Kit according to the manufacturer's instructions.

DETAILED PROTOCOL



STEP 1

RNA fragmentation and dephosphorylation -  *30 minutes*

1.1 Add **1 µl of Fragmentation Buffer (FB)** to the **7 µl of RNA sample**.

1.2 Mix by pipetting up and down at least 5 times.

1.3 Place the sample in a thermal cycler and incubate for 7 minutes at 94°C to fragment the RNA.

***NOTE**: For highly degraded RNA samples (e.g., from FFPE tissues), DO NOT incubate at 94°C after FB addition. Instead, incubate **on ice** for 7 minutes.*

For moderately degraded RNA samples (i.e., fragments > 500 nt), incubate for only 1-5 minutes at 94°C depending on fragment size.

1.4 Spin down the sample in a table top centrifuge and place on ice until further processing.

1.5 Prepare a **Dephosphorylation Master Mix (DMM)** in a RNase-free tube as follows:

Component	Volume
Dephosphorylation Buffer (DB)	2 µl
Dephosphorylation Reagent (DR)	0.3 µl
Total Volume	**2.3 µl**

1.6 Mix by pipetting up and down at least 5 times and use immediately.

1.7 Add **2 µL of DMM** to the sample. Unused **DMM** can be stored at -20°C and thawed up to 3 times.

1.8 Mix by pipetting up and down at least 5 times.

1.9 Incubate for 15 minutes at 37°C on thermal cycler.

1.10 Spin down the sample in a table top centrifuge and place on ice until further processing.

STEP 2

RNA tailing -  *1 hour*

2.1 Prepare a **Tailing Master Mix (TMM)** in a RNase-free tube as follows:

Component	Volume
Tailing Buffer (TB)	1 µl
Tailing Reagent (TR)	0.5 µl
Total Volume	**1.5 µl**

2.2 Mix by pipetting up and down at least 5 times and use immediately.

2.3 Add **1.5 µl of TMM** to the sample.

2.4 Mix by pipetting up and down at least 5 times.

2.5 Spin down the sample briefly (1-2 seconds) in a table top centrifuge.

2.6 Incubate for 40 minutes at 37°C + 20 minutes at 65°C on thermal cycler, and then cool down the sample on ice for 2 minutes.

2.7 Spin down the sample in a table top centrifuge and place on ice until further processing.

STEP 3

Reverse transcription with template switching -  *> 2 hours*

3.1 Add **1 µl of Reverse Transcription Primer UDI (RTP_UDI)** to the sample **on ice**. Choose the right RTP according to your starting RNA amount:

Purified total RNA or enriched RNA	
50 pg – 500 pg	500 pg – 50 ng
RTPM_UDI	RTPH_UDI

3.2 Mix by pipetting up and down until the solution is homogenous.

3.3 Place the sample in a thermal cycler and run the following program:

Temperature	Time
70°C	10 min
70°C - 25°C	0.5°/sec
25°C	2 min hold

3.4 Spin down the sample in a table top centrifuge and place on ice until further processing.

3.5 Take the **Template Switch Oligo UDI (TSO_UDI)** tube out of the freezer and let it thaw on ice.

3.6 Prepare a **Reverse Transcription Master mix (RTMM)** in a RNase-free tube **on ice**, as follows:

Component	Volume
Reverse Transcription Buffer (RTB)	5 µl
Reverse Transcription Reagent (RTR)	1 µl
Total Volume	**6 µl**

3.7 Mix by pipetting up and down at least 5 times and use immediately.

3.8 Add **6 µl of RTMM** to the sample.

3.9 Mix by pipetting up and down until the solution is homogenous.

3.10 Incubate for 15 minutes at 25°C on thermal cycler and then place on ice.

3.11 Add **2 µl of TSO_UDI** to the sample. **TSO** should be stored at -20°C after use.

3.12 Mix by pipetting up and down until the solution is homogeneous.

3.13 Incubate for 120 minutes at 42°C + 10 minutes at 70°C on thermal cycler and then cool down the sample on ice until further processing.

***NOTE**: The Reverse Transcription (120 minutes at 42°C + 10 minutes at 70°C) may be performed overnight. If so, add an additional hold at +4°C once the two steps are completed.*

3.14 Spin down the sample in a table top centrifuge to collect the library at the bottom and proceed with next steps or store the sample at -20°C until further use.

STEP 4

PCR amplification - *1 hour*

4.1 Add **20 µl of D-Plex Primer UDI** (primer pair UDI#) to the sample.

4.2 Add **50 µl of PCR Master Mix (PCRMM)** to the sample.

4.3 Mix by pipetting up and down until solution is homogeneous.

4.4 Determine the number of cycles (n) for PCR amplification according to the RNA input (quantified or estimated after poly(A) capture or rRNA depletion):

Starting amount	n cycles (*)
50 ng	8
10 ng	10
5 ng	12
1 ng	14
100 pg	16
50 pg	17-18

** Depending on the RNA template, quantity and quality of the sample, 1 or 2 PCR cycles can be added to generate more library (but in detriment of the duplicates content). The PCR amplification cycles are only provided here as guidelines and should be adjusted according to your D-Plex RNA-seq experiment.*

4.5 Incubate for PCR amplification according to the following program:

Step	Temperature	Time & Cycles	
1. Initial denaturation	98°C	30 seconds	
2. Denaturation	98°C	15 seconds	*n* cycles
3. Annealing & extension	72°C	60 seconds	
4. Final extension	72°C	10 minutes	
Hold at 4°C or stored at -20°C until further processing			

STEP 5

Library clean-up

5.1 Take the AMPure® XP beads out of the fridge and resuspend them gently on a rotating wheel or orbital shaker at room temperature before use.

5.2 Add **72 µL of AMPure® XP beads (0.8X)** to the **90 µL of the amplified DNA library** from section 5.

5.3 Mix thoroughly by pipetting up and down several times or by vortexing a few seconds.

5.4 Incubate for 5 minutes at room temperature under mild agitation.

5.5 Spin down the sample in a table top centrifuge and place the tube on a magnetic rack for 3 minutes to separate the beads from the supernatant.

5.6 When the solution is clear, carefully remove and discard the supernatant without disturbing the beads.

5.7 While keeping the tube on the magnetic rack, add **200 µL of freshly prepared 80% ethanol** for 30 seconds to wash the beads.

5.8 Carefully remove and discard the supernatant without disturbing the beads.

5.9 Repeat steps 5.7-5.8 once for a total of 2 washes.

5.10 Spin down the sample in a table top centrifuge, place the tube back on the magnetic rack and remove any remaining ethanol.

5.11 Leave the tube open for 2-3 minutes to let the beads pellet dry (do not over-dry the beads as it may result in low recovery, over-dried beads become light brown with cracks) and then remove it from the magnetic rack.

5.12 Add **50 µL of nuclease-free water**.

5.13 Mix slowly by pipetting up and down to resuspend the beads.

5.14 Incubate the beads in water for 2 minutes at room temperature under mild agitation.

5.15 Spin down the sample in a table top centrifuge and place the tube on the magnetic rack for 3 minutes to separate the beads from the supernatant.

5.16 When the solution is clear, carefully collect the **50 µL of supernatant** without taking up any beads.

5.17 Transfer the cleaned-up library in a RNase-free tube.

5.18 Repeat the clean-up procedure (from steps 5.2 to 5.17) with **40 µL AMPure® XP beads (0.8X)** but elute the purified library in **20 µL of nuclease-free water**. If more concentrated library is needed, the elution can be done in 10 µl of nuclease-free water.

5.19 Store the final cleaned-up library at -20°C until further use.

STEP 6

Library QC and quantification

6.1 For quantification, measure the library concentration using **QuBit® dsDNA HS Assay kit** according to the manufacturer's instructions.

6.2 For library size estimation, use the **Bioanalyzer® DNA High Sensitivity assay kit** according to the manufacturer's instructions.

D-Plex Total RNA Construct

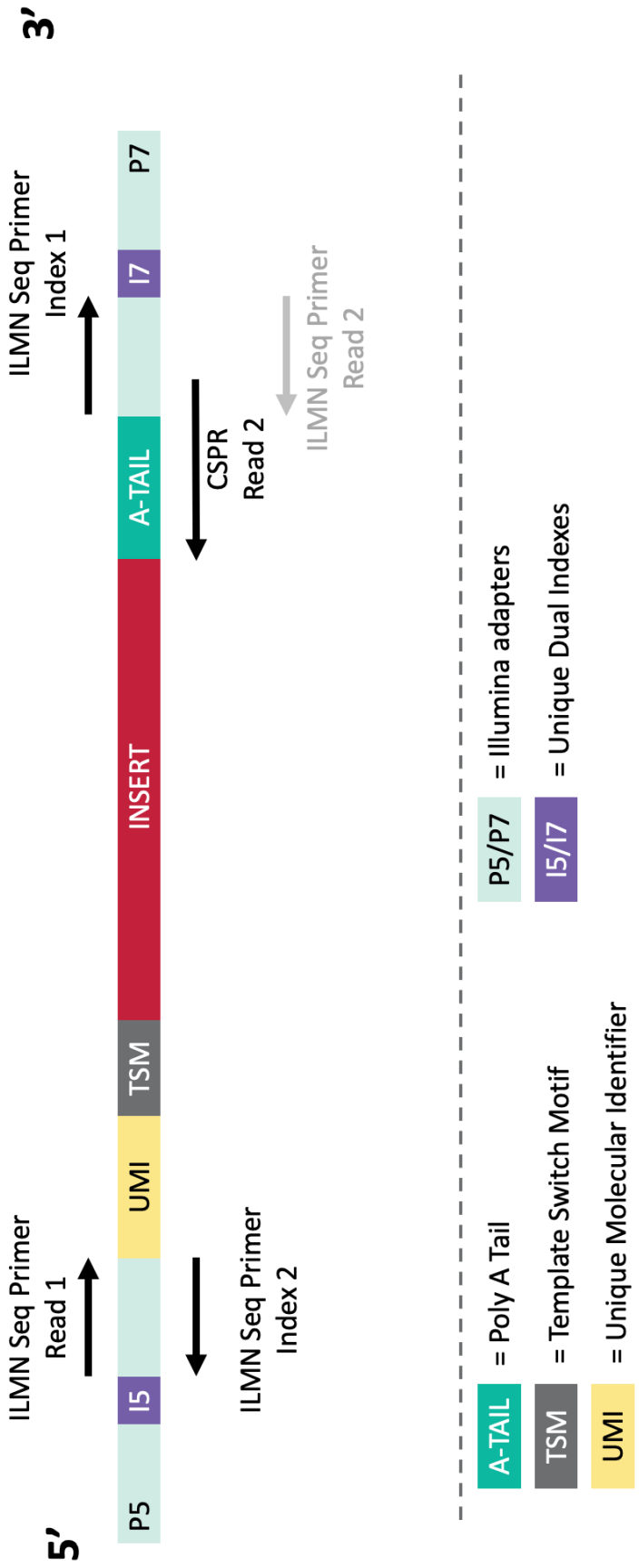


***Figure 5**: The D-Plex Total RNA-seq UDI library construct bears the Truseq (Illumina) HT adapters with unique dual indexes (UDI) and unique molecular identifiers (UMI). Read 1 = UMI – TSM– Insert – A-tail – 3' adapter; Read 2 = Insert – TSM – UMI.*

Sequencing Recommendations

The D-Plex Total RNA-seq library construct bears the TruSeq (Illumina) HT adapters with unique dual indexes (UDI). In case of a multiplexing scenario, it is therefore recommended to submit the D-Plex libraries as TruSeq HT libraries to your sequencing provider.

The complete architecture of the D-Plex Total RNA-seq construct is provided in Figure 5. The empty library size (adapters + template switch + UMI + A tail) is equal to 172bp.

Given the high complexity of whole transcriptome in D-Plex Total RNA-seq libraries, it is advised to sequence one library with a **minimum of 20-25 million reads** in order to sufficiently cover the library content for downstream analysis.

It is recommended to perform the sequencing of the D-Plex libraries in **single-end mode** as the quality of the sequencing data tends to be better than in paired-end. Furthermore, if the interest is mainly gene expression analysis, the current sequencing configurations such as in Illumina's NovaSeq system allow a sufficient (if not total) coverage of the insert by the read length in single end.

The read 1 sequence starts with the UMI and the template switch motif (for a total of 16 bases) that will be removed during the downstream bioinformatics analysis. Therefore, the sequencing read length needs to be adapted. Please follow the recommendations in Table 4.

Example: 1 x 100 bp Single-end sequencing will give (after trimming) a Read 1 sequence length = 84 bp.

Table 4. Requirements for sequencing mode and length

Sequencing	mRNA library	high quality rRNA(-) library	degraded rRNA(-) library
Single-end	1x100bp	1x100bp	1x50bp
phiX %	Follow recommendations of Illumina's sequencing platforms		

However, if paired-end sequencing is needed, we recommend using our Custom Sequencing Primer for Read2 (CSPR2) that has to be ordered separately (C05020001). This custom sequencing primer can be spiked in with the standard Illumina primer for read2. If more information is needed on setting up a paired-end sequencing with the D-Plex libraries, please contact us (customer.support@diagenode.com).

Data Analysis Recommendations

The D-Plex libraries contain special sequences that need particular treatment in order to get the best results out of your datasets. The D-Plex construct holds special sequences namely the UMI, the A-tail, and the template switch motif. This guide will take you through the basic processes of trimming, alignment and counting, complemented with an optional UMI processing, using software tools and settings that we validated. Though naturally other tools and methods can also be used, please pay attention to finding the optimal settings for your experiments. In our example commands, we assume that the necessary software tools have been downloaded from the links provided at the end of the section and are in the PATH.

We recommend the fumi-tools software package for UMI processing. If the UMIs are not of interest, their processing can safely be skipped and the rest of the pipeline will not change. In such a case, the UMI sequences will be only removed from the reads during trimming and will be ignored for the rest of the analysis.

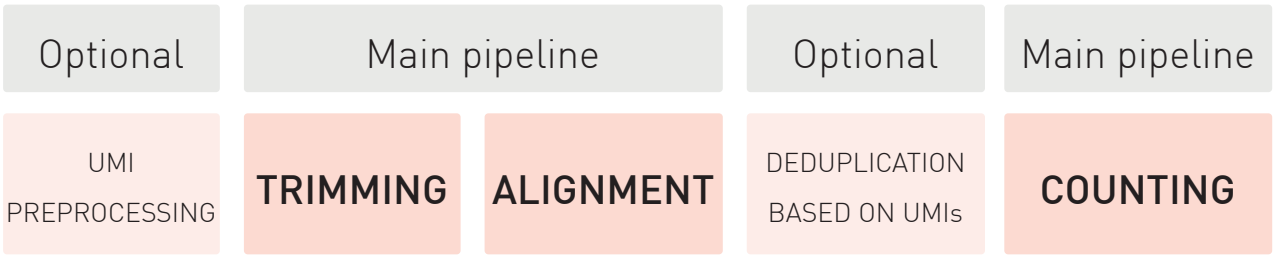


***Figure 6**: D-Plex Total RNA-seq bioinformatics pipeline*

UMI-preprocessing (optional)

To process the UMIs, the first thing to do is to copy the UMI sequence to the read ID. The first 12 bases in the read (from the 5' end) correspond to the UMI sequence. In the fumi-tools package the copy_umi command can copy these bases to the correct lines in the fastq file. The tool expects a fastq file as input, which can be gzip compressed, and it will output a fastq file, which will be either gzip compressed or uncompressed, based on the extension set on the commandline (.gz for compression). Besides the input, the command needs one mandatory option: the length of the UMI

(to copy from the read to the ID line) as input. Optionally, the computational time can be reduced by increasing the number of CPUs to be used via the threads parameter. The example command below uses all the 12 UMI bases (recommended) and 10 threads.

Single-end processing

```
fumi_tools copy_umi --threads 10 --umi-length 12 -i
reads.fastq.gz -o reads_w_umi.fastq.gz
```

Paired-end processing

```
fumi_tools copy_umi --threads 10 --umi-length 12 -i
reads_R1.fastq.gz --input-read2 reads_R2.fastq.gz -o
reads_R1_w_umi.fastq.gz --output-read2 reads_R2_w_
umi.fastq.gz
```

Trimming

Trimming is mandatory for the reads generated with the D-Plex Total RNA-seq kit. The aim of trimming is to remove these artificial sequences typical of D-Plex construct (UMI, A-tail and template switch motif) that will likely hamper downstream analyses.

In our example command below, we use cutadapt package to properly trim the reads and obtain a read set as clean as possible. The commands remove UMI, A-tail/T-tail, template switch motif and adapters sequences from all reads and discard reads which result shorter than 15 bases (these are considered too short to be included in the analysis.). The trimming commands can be applied on the input raw files (reads.fastq.gz) coming directly from the sequencer or alternatively, on the UMI pre-processed files (reads_w_umi.fastq.gz) if the UMIs are of interest.

Single-end processing

```
cutadapt --trim-n --match-read-wildcards -u 16
-n 4 -a AGATCGGAAGAGCACACGTCTG -a AAAAAAAA -a
GAACTCCAGTCAC -e 0.2 --nextseq-trim 20 -m 15 -o
trimmed_reads.fastq.gz raw_reads_w_umi.fastq.gz
```

Paired-end processing

```
cutadapt --trim-n --match-read-wildcards -u
16 -n 4 -a "AGATCGGAAGAGCACACGTCTG;e=0.2" -a
"AAAAAAAA;e=0.2" -a "GAACTCCAGTCAC;e=0.2"' -A
"CCCNNNNNNNNNNNNAGATC;o=21" -G "TTTTT" --nextseq-
trim 20 -m 15:15 -o trimmed_reads.fastq.gz raw_
reads_R1.fastq.gz -p trimmed_reads_R2.fastq.gz raw_
reads_R1_w_umi.fastq.gz raw_reads_R2_w_umi.fastq.gz
```

Alignment

Aligning the trimmed reads needs no special treatment as you can use any aligner that is suitable for mapping RNA-seq reads. First and foremost, we recommend aligning to the genome (instead of the transcriptome). Indeed, D-Plex tends to generate very high-complexity libraries which often include RNAs that are not identified yet, and therefore, would not map to a transcriptome reference consisting of only known transcripts. Of course, in addition to the genome alignment, the mapped reads can be assigned to known transcripts as well for expression analysis of the known genes.

We recommend using the STAR alignment software. To run STAR, you need to provide the following arguments to the program: the input trimmed reads data file; the length at which reads were sequenced minus 1; the folder containing the STAR indexed genome (see STAR user guide for more information on how to prepare the genome); optionally the number of CPUs to be used and the output format.

The example command below shows how to run STAR on the trimmed reads data. We assume we are dealing with a human sample, that reads were sequenced at 50 bases long and that we want to use 10 CPUs.

Please adapt the read length if it is longer than 50 bp and change the hg19 genome if your samples are not human.

Single-end processing

```
STAR --runThreadN 10 --readFilesCommand zcat
--genomeDir /genomes/hg19/ --sjdbGTFfile /genomes/
hg19/hg19.gtf --sjdbOverhang 49 --readFilesIn
trimmed_reads.fastq.gz
```

```
--quantMode TranscriptomeSAM
--quantTranscriptomeBAMcompression -1 --outSAMtype
BAM SortedByCoordinate --outSAMunmapped Within
--outFileNamePrefix ./MySample_
```

Paired-end processing

```
STAR --runThreadN 10 --readFilesCommand zcat
--genomeDir /genomes/hg19/ --sjdbGTFfile /genomes/
hg19/hg19.gtf --sjdbOverhang 49 --readFilesIn
trimmed_reads_R1.fastq.gz trimmed_reads_R2.fastq.gz
```

```
--quantMode TranscriptomeSAM
--quantTranscriptomeBAMcompression -1 --outSAMtype
BAM SortedByCoordinate --outSAMunmapped Within
--outFileNamePrefix ./MySample_
```

Deduplication based on UMIs (optional)

UMI processing enables the distinction between identical reads that are coming from different RNA molecules and identical reads resulting from PCR amplification. (If UMI deduplication is chosen, notice that you should have applied the UMI pre-processing step before trimming). Below we provide the command example to use fumi-tools to remove PCR clones from the alignment files. Computational resources to be used may be changed as desired using threads parameter (number of CPUs) and memory.

To remove duplicates from the genome alignment, no sorting is needed before UMI deduplication. However, to remove duplicates from the transcriptome alignment output, since the alignments are not sorted, please use samtools command provided below before UMI deduplication:

```
samtools sort -@ 10 -o MySample_Aligned.
toTranscriptome.sorted.out.bam MySample_Aligned.
toTranscriptome.out.bam
```

Now you can deduplicate both the genome and the transcriptome alignment with the UMI deduplication command of fumi-tools called dedup:

```
fumi_tools dedup --threads 10 --memory 10G -i
MySample_Aligned.toTranscriptome.sorted.out.bam -o
MySample_deduplicated_transcriptome.bam
```

```
fumi_tools dedup --threads 10 --memory 10G -i
MySample_Aligned.sortedByCoord.out.bam -o MySample_
deduplicated_genome.bam
```

Note that fumi-tools outputs name-sorted bam files, these are ready to be used as input for counting software tools. For another application, you may want to sort the bam files differently (e.g. soring by coordinates).

Counting

The counting, or expression level calculation is the last step of the processing to generate an expression level matrix.

As the D-PLEX Total RNA-seq kit may involve the choice of mRNA capture or rRNA depletion before the library preparation, we recommend having a clear understanding of the scientific question and the goal of the project before proceeding to the choice of the counting method as this will strongly impact the end result.

If the interest is to study the whole transcriptome, we recommend the use of raw counting tools (featureCounts, HTSeq-count) with modified parameters to account for reads overlapping non-coding features and/or introns, if desired.

If the interest is to study the coding part of the transcriptome (exons only), counting tools (f.i. featureCounts, HTSeq-count) may be used for counting exon features and summarizing expression at gene level or alternatively, RSEM may be used to estimate expression at isoform level.

Notice that D-Plex produces forward-stranded data. Stranded libraries have the benefit that reads map to the genome strand where they were originated from. Therefore, when estimating transcript expression, reads aligned to the forward strand should be assigned only to transcript features in the forward strand whereas reads aligned to the reverse strand should be assigned only to transcript features in the reverse strand. For this, make sure you select "stranded-mode" in any tool of choice.

Links for the tools used in the example pipeline

Tool	Link
Website	https://www.diagenode.com/en
Fumi-tools	https://ccb-gitlab.cs.uni-saarland.de/tobias/fumi_tools/releases
Cutadapt	https://github.com/marcelm/cutadapt
STAR	https://github.com/alexdobin/STAR
Samtools	http://www.htslib.org/download
RSEM	http://deweylab.github.io/RSEM
featureCounts	http://subread.sourceforge.net
Htseqcount	https://htseq.readthedocs.io/en/master/count.html

Example of Results

D-Plex Total RNA-seq UDI library construct

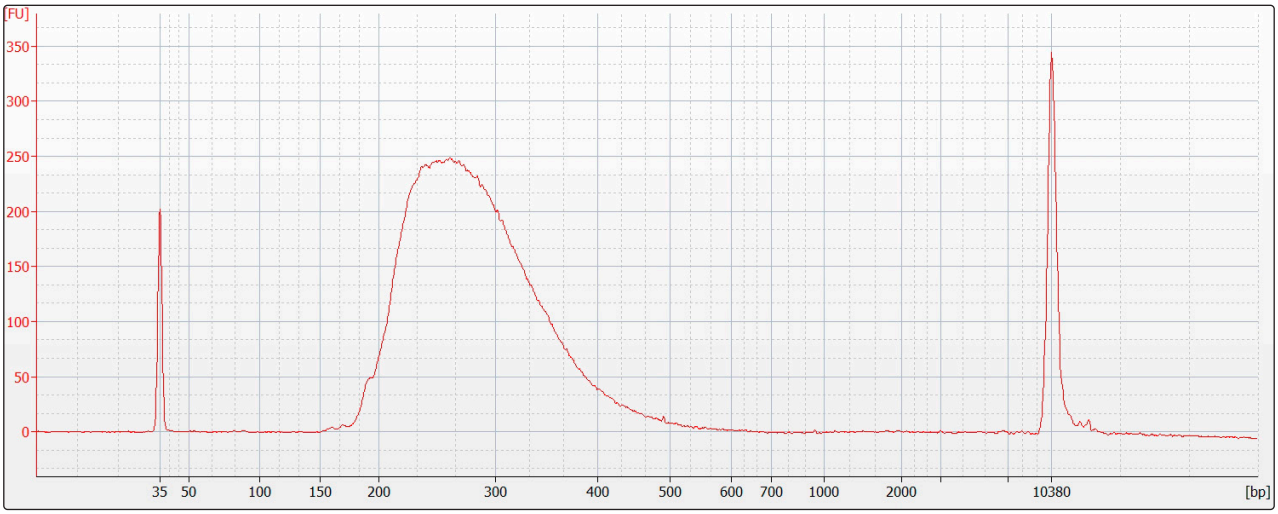


***Figure 7**: DNA electropherogram of a D-Plex Total RNA-seq UDI library construction made from 10 ng of Human Universal Reference RNA before rRNA depletion. 1 ng of the library was loaded on the Bioanalyzer.*

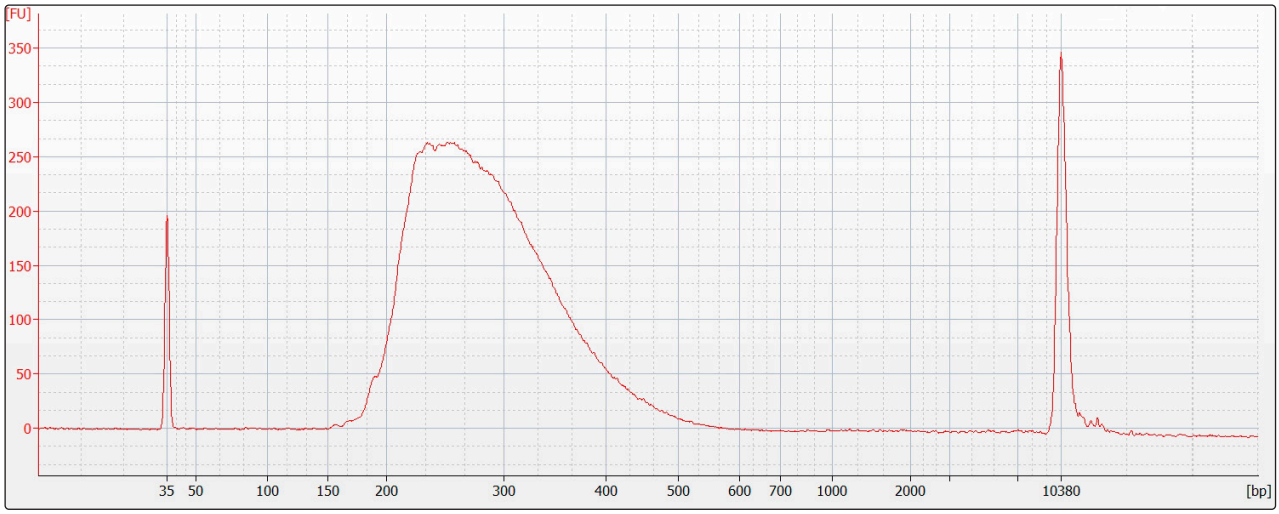


***Figure 8**: DNA electropherogram of a D-Plex Total RNA-seq UDI library construction made from 10 ng of Human FFPE RNA before rRNA depletion. 1 ng of the library was loaded on the Bioanalyzer.*

Related Products

Product	Reference
D-Plex mRNA-seq Kit	C05030033
D-Plex Small RNA-seq Kit	C05030001
D-Plex Unique Dual Indexes Module – Set A	C05030021
D-Plex Unique Dual Indexes Module – Set B	C05030022
D-Plex mRNA Capture Module	C05030032
DiaMag 0.2 mL tube magnetic rack	B04000001
DiaMag 1.5 mL tube magnetic rack	B04000003

FOR RESEARCH USE ONLY.

Not intended for any animal or human therapeutic or diagnostic use.

NOTICE TO PURCHASERLIMITED LICENSE

The information provided herein is owned by Diagenode and/or its affiliates. Subject to the terms and conditions that govern your use of such products and information, Diagenode and/or its affiliates grant you a nonexclusive, nontransferable, non-sublicensable license to use such products and information only in accordance with the manuals and written instructions provided by Diagenode and/or its affiliates. You understand and agree that except as expressly set forth in the terms and conditions governing your use of such products, that no right or license to any patent or other intellectual property owned or licensable by Diagenode and/or its affiliates is conveyed or implied by providing these products. In particular, no right or license is conveyed or implied to use these products in combination with any product not provided or licensed to you by Diagenode and/or its affiliates for such use. Limited Use Label License: Research Use Only The purchase of this product conveys to the purchaser the limited, non-transferable right to use the product only to perform internal research for the sole benefit of the purchaser. No right to resell this product or any of its components is conveyed expressly, by implication, or by estoppel. This product is for internal research purposes only and is not for use in commercial applications of any kind, including, without limitation, quality control and commercial services such as reporting the results of purchaser's activities for a fee orother form of consideration. For information on obtaining additional rights, please contact info@diagenode.com.

TRADEMARKS

The trademarks mentioned herein are the property of Diagenode or their respective owners. Bioanalyzer is a trademark of Agilent Technologies, Inc. Agencourt and AMPure® are registered trademarks of Beckman Coulter, Inc. Illumina® is a registered trademark of Illumina®Inc; Qubit is a registered trademark of Life Technologies Corporation.