Preparation of 6 Whole Transcriptome cDNA libraries and 6 Cappable-Seq libraries for Illumina sequencing

1 Material supplied

Six RNA samples from *Mycobacterium bovis* as indicated in Table 1, delivered on dry ice.

Table 1: Description of the RNA samples

No.	Sample	Conc. (ng/µl)	Amount (µg)	Conc. (ng/µl)	Amount (µg)	Ratio 23S/16S	% after rRNA depletion	Yield Cappable-seq enriched RNA (%)
		Customer-specified		Own measurements (see Fig.1)				
1	MBWT1	351,16	14,05	280	10,6	0,6	3,5	0,2
2	MBWT2	315,82	12,63	273	10,4	0,7	3,3	0,1
3	MBWT3	407,61	16,30	316	12,0	0,7	4,0	0,2
4	MBDSK1	366,51	14,66	295	11,2	0,6	3,2	0,3
5	MBDSK2	288,13	11,53	237	9,2	0,9	3,8	0,4
6	MBDSK3	341,13	13,65	251	9,8	0,6	4,0	0,2

2 Analysis of total RNA and rRNA depletion

The total RNA samples were examined by capillary electrophoresis (Fig. 1). From aliquots of the total RNAs, ribosomal RNA molecules were depleted using the in-house developed depletion probes. Recovery rates are shown in Table 1.

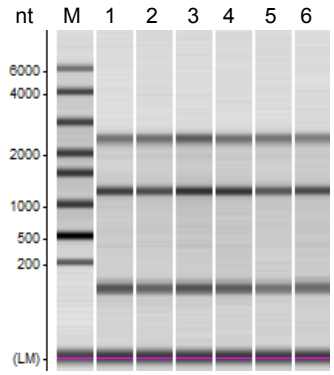

Figure 1: Analysis of the total RNA samples on a Shimadzu MultiNA microchip electrophoresis system. M = RNA marker

3 cDNA synthesis

3.1 Whole Transcriptome cDNA

The ribodepleted RNA samples were first fragmented using ultrasound (4 pulses of 30 s each at 4°C). Then, an oligonucleotide adapter was ligated to the 3' end of the RNA molecules. First-strand cDNA synthesis was performed using M-MLV reverse transcriptase and the 3’ adapter as primer. The first-strand cDNA was purified and the 5' Illumina TruSeq sequencing adapter was ligated to the 3' end of the antisense cDNA. The resulting cDNA was PCR-amplified to about 10-20 ng/µl using a high fidelity DNA polymerase (cycle numbers are indicated in Table 2). The TruSeq barcode sequences which are part of the 5' and 3' TruSeq sequencing adapters are included in Table 2. The cDNA was purified using the Agencourt AMPure XP kit (Beckman Coulter Genomics) and was analyzed by capillary electrophoresis (Fig. 2).

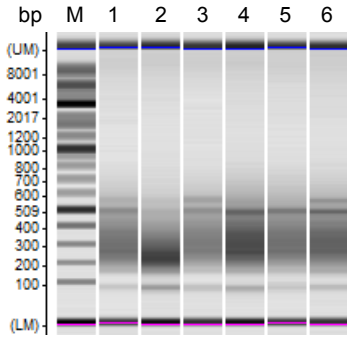

Figure 2: Analysis of the PCR-amplified Whole Transcriptome samples on a Shimadzu MultiNA microchip electrophoresis system. M = 100 bp ladder

3.2 cDNA from Cappable-seq enriched RNA

Aliquots from the total RNA samples were enriched for primary transcripts carrying 5'-tri-phosphate ends. The following steps were carried out:

- Capping of the 5' triphosphorylated RNA with 3'- desthiobiotin-TEG-guanosine 5' triphosphate (DTBGTP) (NEB) using the vaccinia capping enzyme (VCE) (NEB) for reversible binding of biotinylated RNA species to streptavidin
- Capturing of biotinylated RNA species on streptavidin beads and elution to obtain the 5' fragment of the primary transcripts

Yields of Cappable-seq enriched RNAs are shown in Table 1.

In order to be able to identify contamination of the Cappable enriched RNAs with processed transcripts carrying 5'-monophosphate ends, our tagRNA-seq method was used for cDNA synthesis. The Cappable-seq enriched RNAs were first poly(A)-tailed using poly(A) polymerase. Then, the 5' Illumina TruSeq sequencing adapters, which carry sequence tags ATTACTCG and TCCGGAGA (in a proportion of 50% of each adapter), were ligated to the 5' mono-phosphate groups (5'P) of processsed transcripts. The samples were then treated with CapClip Acid Pyrophosphatase (Cellscript) in order to convert 5' triphosphate (5'PPP) structures into 5' monophosphate ends. To the newly formed 5'P groups the 5' Illumina TruSeq sequencing adapters, which carry sequence tag CGCTCATT and GAGATTCC (50% each), were ligated. First-strand cDNA synthesis was performed using an oligo(dT)-adapter primer and the M-MLV reverse transcriptase. The resulting cDNAs were PCR-amplified to about 10-20 ng/µl using a high fidelity DNA polymerase (Fig. 3 A; cycle numbers are indicated in Table 2).

For Illumina sequencing, 100 – 300 bp long 5' fragments were isolated from the full-length cDNAs. For this purpose the cDNA preparations fragmented and the 5'-cDNA fragments were then bound to streptavidin magnetic beads. The bound cDNAs were blunted and the 3' Illumina sequencing adapter was ligated to the 3' ends of the cDNA fragments. The bead bound cDNAs were finally PCR-amplified (see Fig. 3 B). PCR cycles performed and the barcode sequences, which are attached to the 5' and 3' ends of the cDNAs, are included in Table 2.

Table 2. Properties of cDNA samples

No.	Sample	Method	5' barcode	3' barcode	PCR cycles	PCR cycles (5')
1	MBWT1	WT	ACGTCCTG	TCCGGAGA	14	-
2	MBWT2	WT	GTCAGTAC	TCCGGAGA	14	-
3	MBWT3	WT	ATAGAGAG	TCCGGAGA	14	-
4	MBDSK1	WT	AGAGGATA	TCCGGAGA	14	-
5	MBDSK2	WT	CTCCTTAC	TCCGGAGA	14	-
6	MBDSK3	WT	TATGCAGT	TCCGGAGA	14	-
1C	MBWT1	Cappable-seq	AGGCTATA	ATTACTCG	15	6
2C	MBWT2	Cappable-seq	GCCTCTAT	ATTACTCG	15	6
3C	MBWT3	Cappable-seq	AGGATAGG	ATTACTCG	15	6
4C	MBDSK1	Cappable-seq	TCAGAGCC	ATTACTCG	15	6
5C	MBDSK2	Cappable-seq	CTTCGCCT	ATTACTCG	15	6
6C	MBDSK3	Cappable-seq	TAAGATTA	ATTACTCG	15	6

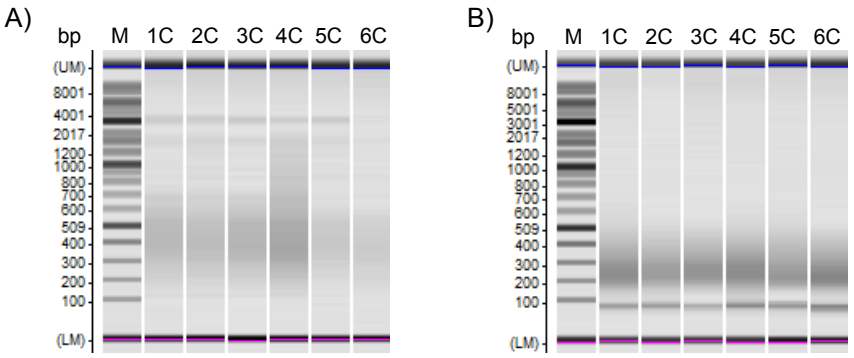

Figure 3: Analysis of the PCR-amplified Cappable-seq full-length cDNA (A) and the PCR-amplified 5' fragment cDNAs (B) on a Shimadzu MultiNA microchip electrophoresis system. M = 100 bp ladder

4 Pool generation and size fractionation

For Illumina NextSeq sequencing, the samples were pooled in approximately equimolar amounts. The cDNA pool was size fractionated in the size range of 200 - 550 bp using a preparative agarose gel. An aliquot of the size fractionated pool was analyzed by capillary electrophoresis (Fig. 4).

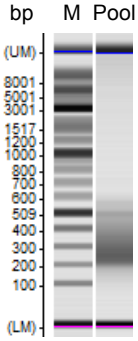

Figure 4: Analysis of the size-fractionated cDNA pool on a Shimadzu MultiNA microchip electrophoresis system. M = 100 bp ladder

5 Sample description

The primers used for PCR amplification were designed for TruSeq sequencing according to the instructions of Illumina.

The following adapter sequences flank the cDNA inserts:

TruSeq_Sense_primer i5 Barcode
5'-AATGATACGGCGACCACCGAGATCTACAC-NNNNNNNN-ACACTCTTTCCCTACACGACGCTCTTCCGATCT-3'

TruSeq_Antisense_primer i7 Barcode
5'-CAAGCAGAAGACGGCATACGAGAT-NNNNNNNN-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'

The combined length of the flanking sequences is 136 bases.

In addition, the Cappable-seq cDNA carries the following 5' tag sequences:

5' ends of cDNA from processed RNAs (before CapClip treatment)
(linker 5-15 nt) (linker 5-15 nt)
NNATTACTCGNN-cDNA insert and NNTCCGGAGANN-cDNA insert

5' ends of cDNA from primary transcripts (after CapClip treatment)
(linker 5-15 nt) (linker 5-15 nt)
NNCGCTCATTNN-cDNA insert and NNGAGATTCCNN-cDNA insert

6 Illumina sequencing

The cDNA pool was sequenced on an Illumina NextSeq 500 system using 75 bp read length.