Please cite this article as: Jun *et. al.*, (2016). Gene Expression Analysis of Sorted Cells by RNA-seq in *Drosophila* Intestine, *Bio-protocol* 6 (24): e2079. DOI: 10.21769/BioProtoc.2079.

bio-protocol www.bio-protocol.org/e2079 Vol 6, Iss 24, Dec 20, 2016 DOI:10.21769/BioProtoc.2079

Gene Expression Analysis of Sorted Cells by RNA-seq in *Drosophila* Intestine

Jun Chen, Jia Li, Huaiwei Huang and Rongwen Xi*

National Institute of Biological Sciences, Beijing, China
*For correspondence: xirongwen@nibs.ac.cn

[Abstract] RNA sequencing (RNA-seq) has become a popular method for profiling gene expression. Among many applications, one common purpose is to identify differentially expressed genes and pathways in different biological or pathological conditions. This protocol provides detailed procedure for RNA-seq analysis of ~250,000 sorted *Drosophila* intestinal cells (Chen *et al.*, 2016), in which RNA amplification is not required.

Keywords: RNA-seq, FACS, *Drosophila* intestine, Progenitor cells, Transcriptome analysis

[Background] Transcriptome analysis by RNA-seq has become a popular method for the identification of differentially expressed genes and pathways under different biological or pathological conditions. For samples that yield low mRNA levels, RNA or cDNA amplification was commonly performed before deep-sequencing (Dutta *et al.*, 2015). However, this procedure could potentially omit important candidates that are expressed in low abundance. Here we provide a detailed procedure for RNA-seq analysis of sorted *Drosophila* gut cells in which RNA amplification is not required.

Materials and Reagents

A. Isolation of intestinal progenitor cells by FACS
   1. Microcentrifuge tube (Corning, Axygen®, catalog number: MCT-150-C)
   2. 40 μm filters
   3. 70 μm filters (Corning, Falcon®, catalog number: 352350)
   4. The fly strains carrying Esg-GFP fluorescent marker
   5. Elastase (Sigma-Aldrich, catalog number: E0258)
   6. RNAiso Plus (Takara Bio, catalog number: 9108/9109)
   7. NaCl
   8. KCl
   9. Na_2HPO_4
   10. KH_2PO_4
   11. Diethyl pyrocarbonate (DEPC) (Sigma-Aldrich, catalog number: D5758)
   12. 1x DEPC-PBS (see Recipe)
   13. Elastase solution (see Recipe)

Please cite this article as: Jun *et. al.*, (2016). Gene Expression Analysis of Sorted Cells by RNA-seq in *Drosophila* Intestine, *Bio-protocol* 6 (24): e2079. DOI: 10.21769/BioProtoc.2079.

B. RNA isolation
1. Directzol RNA MiniPrep Kit (ZYMO RESEARCH, catalog number: R2050)
2. 95-100% ethanol

C. Library construction
1. Agilent RNA 6000 Pico Kit (Agilent Technologies, catalog number: 5067-1513)
2. Dynabeads® mRNA DIRECT™ Purification Kit (Thermo Fisher Scientific, Ambion™, catalog number: 61011)
3. Hexadeoxyribonucleotide mixture, pd(N)6 (Takara Bio, catalog number: 3801)
4. dNTP mixture (Takara Bio, catalog number: 4019)
5. DTT (provided by SuperScript® II reverse transcriptase)
6. Recombinant RNasin® ribonuclease inhibitor (Promega, catalog number: N2511)
7. SuperScript® II reverse transcriptase (Thermo Fisher Scientific, Invitrogen™, catalog number: 18064014)
8. Second-strand buffer (Thermo Fisher Scientific, Invitrogen™, catalog number: 10812014)
9. DNA polymerase I (Takara Bio, catalog number: 2130A)
10. RNase H (New England Biolabs, catalog number: M0297S)
11. Agencourt AMPure XP Kit (5 ml) (Beckman Coulter, catalog number: A63880)
12. NEBNext® DNA library prep master mix set for Illumina® (New England Biolabs, catalog number: E6040L)
13. NEBNext® multiplex oligos for Illumina® (New England Biolabs, catalog number: E7335S/E7500S)
14. Agilent High Sensitivity DNA Kit (Agilent Technologies, catalog number: 5067-4626)
15. Qubit® dsDNA HS Assay Kit (Thermo Fisher Scientific, Invitrogen™, catalog number: Q32851)
16. Illumina Library Quantification Kit (Kapa Biosystems, catalog number: KK4824)

Equipment

A. Isolation of intestinal progenitor cells by FACS
1. Dissecting microscope with zoom and dual goosenecks to supply oblique illumination; CO_2 equipped fly sorting station (Leica, model: MZ16; custom fabrication)
2. Forceps, Dumont #5 (Fine Science Tools, catalog number: 11252-30)
3. Dissecting dish (Thermo Fisher Scientific, Fisher Scientific, catalog number: 21-379)
4. FACS Aria II sorter (BD)

B. RNA isolation
1. Vortex

Please cite this article as: Jun *et. al.*, (2016). Gene Expression Analysis of Sorted Cells by RNA-seq in *Drosophila* Intestine, *Bio-protocol* 6 (24): e2079. DOI: 10.21769/BioProtoc.2079.

C. Library construction
1. DynaMag™-2 magnet (Thermo Fisher Scientific, catalog number: 12321D)
2. Agencourt AMPure XP 5 ml Kit (Beckman Coulter, catalog number: A63880)
3. PCR thermal cycler
4. Qubit® 3.0 Fluorometer (Thermo Fisher Scientific, Invitrogen™, model: Qubit® 3.0 Fluorometer)
5. Agilent 2100 Bioanalyzer (Agilent Technologies, model: 2100 Bioanalyzer)
6. Applied Biosystems™ 7500 Fast & 7500 real-time PCR system (Thermo Fisher Scientific, model: 7500 Fast & 7500 real-time PCR system)
7. Hiseq-2500 sequencing system (Illumina, model: Hiseq-2500)

Software

1. Mapping and analysis of Illumina reads for transcriptome
   a. Bowtie (http://bowtie-bio.sourceforge.net/index.shtml, Langmead *et al.*, 2009)
   b. TopHat (http://ccb.jhu.edu/software/tophat/index.shtml, Trapnell *et al.*, 2009)
   c. Cufflinks (http://cole-trapnell-lab.github.io/cufflinks/, Roberts *et al.*, 2011)
   d. CASAVA (http://support.illumina.com.cn/sequencing/sequencing_software/casava.html, Illumina)

Procedure

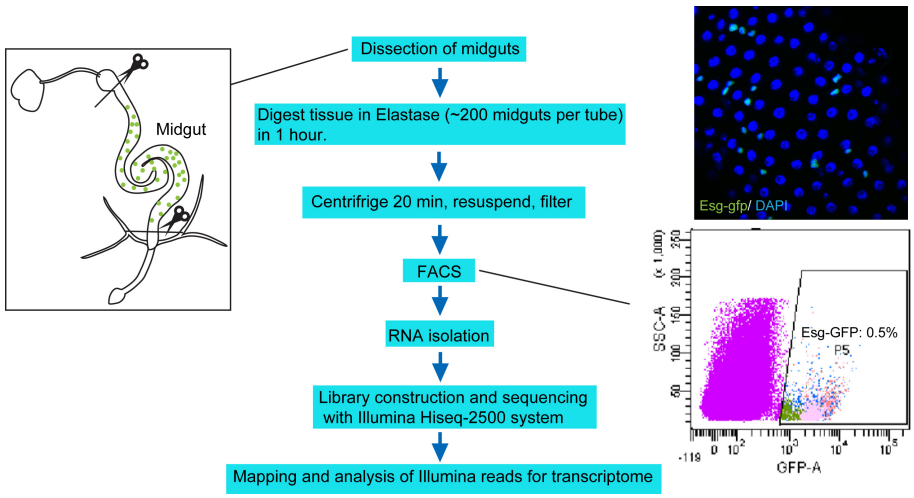

Figure 1. Protocol overview. Under RNase-free environment, the midguts with foregut and hindgut portion removed (left panel) are digested with Elastase for 1 h. The dissociated tissues are centrifuged, resuspended in DEPC-PBS, filtered, and then sorted through FACS. About 250,000 sorted cells are collected to harvest total RNA, which is then used for library construction and sequencing with Illumina Hiseq-2500 system. The upper right panel shows a confocal image of midgut epithelium with Esg-GFP expression. The bottom right panel highlights the Esg-GFP$^+$ cell population measured by FACS.

Please cite this article as: Jun *et. al.*, (2016). Gene Expression Analysis of Sorted Cells by RNA-seq in *Drosophila* Intestine, *Bio-protocol* 6 (24): e2079. DOI: 10.21769/BioProtoc.2079.

A. Isolation of intestinal progenitor cells by FACS (Figure 1)

1. The fly strains carrying Esg-GFP fluorescent markers are used to sort intestinal progenitor cell population (including intestinal stem cells and enteroblasts), midguts from w^{1118} strain are used to set fluorescence gate. To achieve 250K Esg-GFP^{+} cells, a minimum of 1,000 midguts is required.

   Note: It is important to be consistent with age, gender and culture conditions among different biological samples. Here we take expression analysis of Sox21a mutant flies for example. Sox21a mutants carrying Esg-GFP markers served as the mutant sample, with Esg-GFP wild-type flies as WT Ctrl. Based on the observation that overproliferation phenotype is displayed after 10 days in 90% sox21a mutant intestines, we analyzed flies at 10 days old for both the mutants and WT control. For both groups, we chose female flies and cultured them with identical food (standard fly food plus yeast paste).

2. The following steps are performed under RNase-free conditions. Forceps and the dissecting pads are pre-washed with DEPC-PBS solution. The prepared females are ice-anesthetized for dissection. Intestines are dissected and the foregut and hindgut parts are removed by forceps (Figure 2). The midguts are then immediately put in DEPC-PBS (see Recipes) solution on ice.

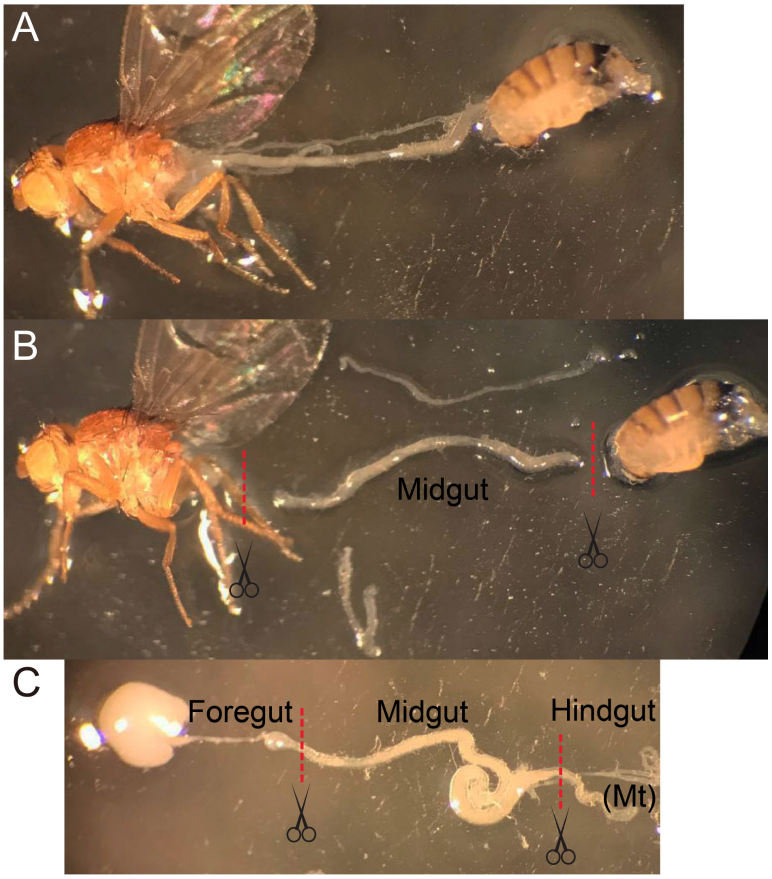

Figure 2. The midgut dissection procedure. A. Use forceps to gently hold a fly and tear the abdomen at the boundary of thorax/abdomen, pull the abdomen away from the anterior part without touching the gut. B and C. Before the gut is fully stretched, cut the gut at the boundaries

Please cite this article as: Jun *et. al.*, (2016). Gene Expression Analysis of Sorted Cells by RNA-seq in *Drosophila* Intestine, *Bio-protocol* 6 (24): e2079. DOI: 10.21769/BioProtoc.2079.

between foregut and hindgut. Remove the appendix if Malpighian tubule (Mt) or ovarium is attached to the midgut.

3. Incubate 100-200 guts with 1 mg/ml Elastase (see Recipes) in about 1 ml DEPC-PBS per microcentrifuge tube for 1 h at 25 °C until the midguts are largely dissociated. Softly mix the sample every 15 min by pipetting and inverting several times.
4. Dissociated samples are pelleted at 400 *x g* for 20 min, and resuspended in a microcentrifuge tube with 0.5 ml DEPC-PBS. The suspension was filtered with 40- or 70-μm filters (Corning), by touching pipette tips on the top of the filter so that the cells can go through filters. Wash the tubes and the filters with 0.5 ml DEPC-PBS, and also collect the filtered suspension in the filtered cells. The filtered cells are then sorted using a FACS Aria II sorter (BD Biosciences) and collected in 0.5 ml DEPC-PBS.
   Note: 40-μm filters are recommended for pure isolation of small diploid cells, such as intestinal progenitor cells or enteroendocrine cells. It's recommended to double check the sorted cells based on cell morphology and expression of the fluorescent marker.
5. The sorted cells are pelleted at 400 *x g* for 20 min and preserved in 0.2-0.8 ml RNAiso Plus (Takara Bio) reagents, which can be stored at -80 °C within 6 months until RNA isolation. For each of the biological replicates, at least 200,000 sorted cells are collected to harvest total RNA. For each genotype group, at least 3 biological replicates are prepared.

B. RNA isolation
   1. Add equal volume of > 95% ethanol directly to the homogenate cell sample. Mix thoroughly by vortexing.
   2. Follow the Directzol RNA MiniPrep Kit to harvest total RNA of each sample. DNase I treatment is unnecessary, as polyA RNAs will be selectively purified during cDNA library construction. The RNA sample eluted in DNase/RNase-free water should be used immediately for sequencing library construction or stored at -80 °C for up to 3 months.

C. Library construction
   1. Qualify total RNA on Agilent 2100 Bioanalyzer using Agilent RNA 6000 Pico Kit. Purify polyA RNA from 50 ng total RNA using Oligo-dT Dynabeads according to manufacturer's protocol (Dynabeads® mRNA DIRECT™ Purification Kit) and elute the polyA RNA in 12 μl RNase free water.
   2. RNA fragmentation

PolyA RNA	12 μl
5x first strand buffer	6 μl

   Incubate at 95 °C 5 min and chill on ice

Please cite this article as: Jun *et. al.*, (2016). Gene Expression Analysis of Sorted Cells by RNA-seq in *Drosophila* Intestine, *Bio-protocol* 6 (24): e2079. DOI: 10.21769/BioProtoc.2079.

3. Reverse transcription

   Add 1.5 µl Hexadeoxyribonucleotide mixture (50 ng/µl) to each sample, incubate at 65 °C for 5 mins and chill on ice.

   Prepare the following RT mix on ice (1x):

dNTP mixture (10 mM)	1.5 µl
DTT (0.1 M)	3 µl
Ribonuclease inhibitor (40 U/µl)	1 µl
SuperScript II Reverse Transcriptase (200 U/µl)	1 µl
Nuclease-free water	4µl

   Add 10.5 µl of the RT mix to each sample, mix and spin. Transfer to thermocycler:

   25 °C 10 min

   42 °C 50 min

   70 °C 15 min

   4 °C hold

4. Synthesis second strand of cDNA

First strand cDNA	30 µl
dNTP mixture (10 mM)	2 µl
Second strand buffer (5x)	10 µl
RNase H (5 U/µl)	1 µl
DNA polymerase I (5 U/µl)	3 µl
Nuclease-free water	4 µl

   16 °C 2.5 h

   4 °C hold

   Purify cDNA using 1.8x AMPure XP beads, and elute in Nuclease-free water.

5. Construct cDNA library using NEBNext® DNA library prep master mix set and NEBNext® multiplex oligos.
6. Qualify library on Agilent 2100 Bioanalyzer using Agilent High Sensitivity DNA Kit. Quantify library using Qubit® dsDNA HS Assay Kit and Illumina Library Quantification Kit according to manufacturer's protocol. Run library on the Illumina Hiseq-2500 sequencing system.

Data analysis

A. Mapping and analysis of Illumina reads for transcriptome
   1. Illumina Casava1.8 software used for basecalling.
   2. *Drosophila melanogaster* genome sequences and bowtie index can be downloaded from Tophat website (Figure 3).

Please cite this article as: Jun *et. al.*, (2016). Gene Expression Analysis of Sorted Cells by RNA-seq in *Drosophila* Intestine, *Bio-protocol* 6 (24): e2079. DOI: 10.21769/BioProtoc.2079.

bio-protocol www.bio-protocol.org/e2079 Vol 6, Iss 24, Dec 20, 2016 DOI:10.21769/BioProtoc.2079

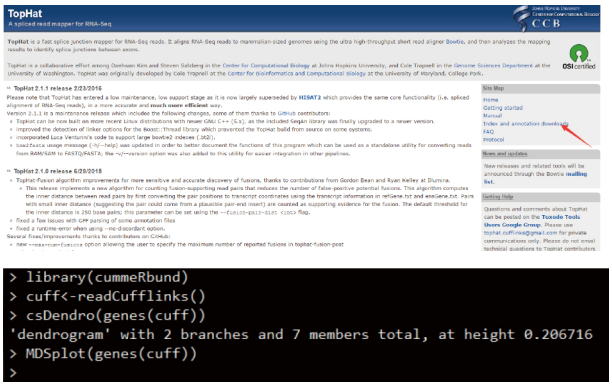

Figure 3. Screenshots of the Tophat webpage

3. Single-end reads are mapped to the *Drosophila melanogaster* genome (Release 5) using TopHat (v2.0.10). For our analysis, we allow up to 2 mismatches and aligned with command 'tophat --bowtie1 –N 2 --no-coverage-search –G genes.gtf genome sample.fastq', where genes.gtf is the structure of genes and sample.fastq is the raw sequencing reads of a sample.
4. Cuffquant is used for quantifying gene and transcript expression levels for a sample BAM file with command 'cuffquant -o outputdir genes.gtf -p 15 sample_aligned/accepted_hits.bam'. This step generates .cxb files which contain gene and transcript expression levels.
5. We use cuffnorm for computing normalized expression values (FPKM) of genes of all samples. csDendro and MDSplot are used for provide insight into the relationships between replicates and different conditions. This step makes sure the replicates samples are consistent and can be used for subsequent analysis (Figure 4).
6. Differentially expressed genes are identified by cuffdiff and significant differentially expression genes were filtered with *P* value ≤ 0.05, and fold change ≥ 2.

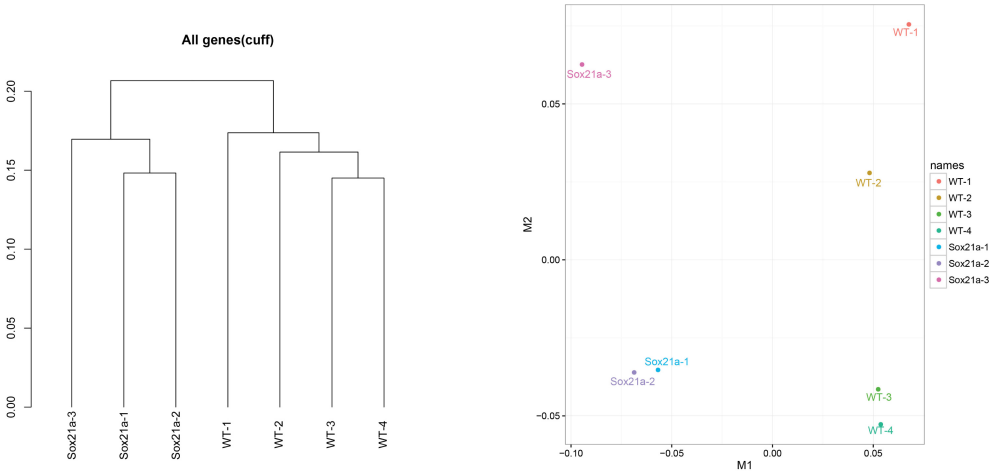

Figure 4. csDendro (left) and MDSplot (right) are used to evaluate the relationships among experimental and control groups. Here the example shows 3 replicates of *sox21a* mutant samples (Sox21a-1, -2, -3) and 4 replicates of WT control samples (WT-1, -2, -3, -4).

Please cite this article as: Jun *et. al.*, (2016). Gene Expression Analysis of Sorted Cells by RNA-seq in *Drosophila* Intestine, *Bio-protocol* 6 (24): e2079. DOI: 10.21769/BioProtoc.2079.

Recipes

1. 1x DEPC-PBS
   1x PBS solution (pH 7.2-7.4) contains 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4 and 1.4 mM KH_2PO_4
   0.1% final solution of diethyl pyrocarbonate (DEPC) is added to 1x PBS
   Mix well by shaking and leave overnight in a fume hood at room temperature
   The solution can be stored at room temperature up to a year in RNase-free conditions
2. Elastase solution
   Dissolve elastase in 1x DEPC-PBS buffer at a final concentration of 1 mg/ml

Acknowledgments

Sorting of intestinal progenitor cells was performed according to the method described previously (Dutta *et al.*, 2015) with some modifications. This work was supported by National Basic Science 973 grant (2014CB850002) from the Chinese Ministry of Science and Technology.

References

1. Chen, J., Xu, N., Huang, H., Cai, T. and Xi, R. (2016). A feedback amplification loop between stem cells and their progeny promotes tissue regeneration and tumorigenesis. *Elife* 5. pii: e14330
2. Dutta, D., Buchon, N., Xiang, J. and Edgar, B. A. (2015). Regional cell specific RNA expression profiling of FACS isolated *Drosophila* intestinal cell populations. *Curr Protoc Stem Cell Biol* 34: 2F 2 1-14.
3. Langmead, B., Trapnell, C., Pop, M. and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10(3): R25.
4. Trapnell, C., Pachter, L. and Salzberg, S. L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25(9): 1105-1111.
5. Roberts, A., Trapnell, C., Donaghey, J., Rinn, J. L. and Pachter, L. (2011). Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol* 12(3): R22.