RNA sequencing of FACS-sorted immune cell populations from zebrafish infection models to identify cell specific responses to intracellular pathogens

Rougeot, J.J.Y.; Zakrzewska, A.; Kanwal, Z.; Jansen, H.J.; Spaink, H.P.; Meijer, A.H.

Chapter 15

RNA Sequencing of FACS-Sorted Immune Cell Populations from Zebrafish Infection Models to Identify Cell Specific Responses to Intracellular Pathogens

Julien Rougeot, Ania Zakrzewska, Zakia Kanwal, Hans J. Jansen, Herman P. Spaink, and Annemarie H. Meijer

Abstract

The zebrafish (*Danio rerio*) is increasingly used as a model for studying infectious diseases. This nonmammalian vertebrate host, which is transparent at the early life stages, is especially attractive for live imaging of interactions between pathogens and host cells. A number of useful fluorescent reporter lines have recently been developed and significant advances in RNA sequencing technology have been made, which now make it possible to apply the zebrafish model for investigating changes in transcriptional activity of specific immune cell types during the course of an infection process.

Here we describe how to sequence RNA extracted from fluorescently labeled macrophages obtained by cell-sorting of 5-day-old zebrafish larvae of the transgenic *Tg(mpeg1:Gal4-VP16);Tg(UAS-E1b:Kaede)* line. This technique showed reproducible results and allowed to detect specific expression of macrophage markers in the mpeg1 positive cell population, whereas no markers specific for neutrophils or lymphoid cells were detected. This protocol has been also successfully extended to other immune cell types as well as cells infected by *Mycobacterium marinum*.

Key words RNA sequencing, FACS, Immune cells, Zebrafish larvae dissociation, Transcriptome analysis

1 Introduction

Infection is associated with complex changes in gene expression patterns of both host and pathogen [1]. An insight into these transcriptional programs can help in identifying new virulence determinants and mechanisms of host defense. The development of genome-wide RNA sequencing (RNAseq) over the last 5 years has revolutionized our approach of transcriptomics [2]. RNAseq consists of a massively parallel sequencing of cDNA obtained from a RNA sample. The millions of sequences obtained (called reads) are then mapped onto a reference sequence in order to assess the presence and the expression level of a transcript in the sample.

Annette C. Vergunst and David O'Callaghan (eds.), *Host-Bacteria Interactions: Methods and Protocols*,
Methods in Molecular Biology, vol. 1197, DOI 10.1007/978-1-4939-1261-2_15,

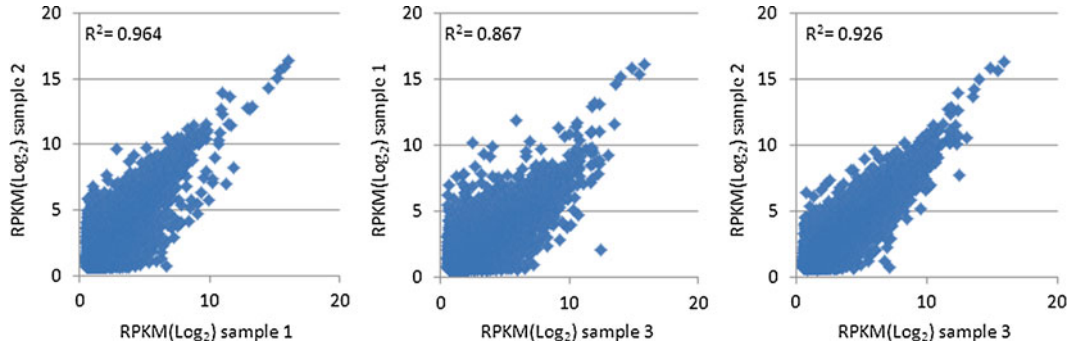

Fig. 1 Reproducibility between three biological replicates. The square of the Pearson correlation coefficient (R^2) was calculated between each replicate sample obtained from fluorescence positive cells based on RPKM (read count per kilobase per million mapped reads) values of genes detected in both replicates (i.e., RPKM >0.5 in both samples). A scatter plot is used to represent gene counts (i.e., Log_2 RPKM values) for replicate samples. Gene count reproducibility was remarkably high considering the numerous steps required to obtain RNAseq libraries from FACS sorted zebrafish larvae. Nevertheless, variation between samples (for example samples 1 and 3 show lower reproducibility than the other combinations) indicates that at least three replicates are required to obtain good statistical results

cDNA read lengths of over a hundred nucleotides can now routinely be obtained with the use of paired-end technology to link the ends of short cDNA fragments [3]. Thus, RNAseq has been described as a powerful method to characterize transcriptional landscapes and discover novel transcripts or alternative splice forms [4]. RNAseq has also proved to be an accurate method for quantitative analysis of differential gene expression [5]. With the development of efficient cDNA synthesis and library preparation protocols, it is now possible to determine the transcriptome of very small populations of cells or even of single cells [6–9].

The zebrafish has recently emerged as a nonmammalian vertebrate model to study host–pathogen interactions, providing many versatile tools for genetics and intravital imaging [10–13]. Zebrafish infection models have been developed for a number of intracellular pathogens, such as *Burkholderia*, *Listeria*, *Mycobacterium*, *Salmonella*, and *Staphylococcus* species [14–18]. Tag-based and full mRNA sequencing analyses have already been used to study the transcriptome of adult zebrafish or embryos in response to pathogen challenge [19–23]. However, until now these studies were limited to determining the immune response at whole organism or organ level. The development of fluorescent reporter lines for different immune cell types [24–28], together with the latest advances in RNAseq technologies, has now made the sequencing of specific immune cell populations in zebrafish feasible.

Here we describe a protocol to sequence the transcriptome of macrophages obtained by cell-sorting of 5 day-old *Tg(mpeg1:Gal4-VP16);Tg(UAS-E1b:Kaede)* transgenic zebrafish larvae [24]. With this protocol, we succeeded in obtaining three reproducible replicates of the fluorescence positive macrophage transcriptome (Fig. 1).

Differential analysis of gene expression between fluorescence positive and fluorescence negative cells resulted in the detection of several known macrophage markers (e.g., *mpeg1*, *mhc2dab*, *mfap4*, *csf1ra*, *marco*, *irf8*), general myeloid markers (e.g., *spi1*), and pan-leukocytic markers (e.g., *coro1a*, *ptprc*, *ptpn6*), whereas no markers specific for neutrophils (e.g., *mpx*) or lymphoid cells (e.g., *lck*, *rag1*, *rag2*) were detected in the enriched pool of fluorescence positive cells.

With this protocol, we have also succeeded in determining the transcriptome of neutrophils and pre-lymphoid cells obtained by cell-sorting of respectively 5 day-old Tg(*mpx:eGFP*) [27] and Tg(*lck:eGFP*) transgenic larvae [26] (data not shown). Furthermore, this protocol has successfully been used to sort and sequence the transcriptome of infected cells obtained from AB/TL larvae infected with *Mycobacterium marinum* expressing a mCherry fluorescence marker (unpublished data) and it is currently being used in our laboratory to investigate the transcriptional reprogramming of macrophages during different stages of *M. marinum* infection. The next challenge in RNAseq analyses for host–pathogen interaction studies is the sequencing of both host cell and pathogen transcriptomes simultaneously. The development of the third generation sequencing platforms could make this so-called dual RNAseq feasible within the next year [1]. In this chapter we describe how to dissociate larvae by trypsin treatment and sort fluorescently labeled cells from the resulting single cell suspension by FACS. Subsequently, we explain how RNA extracted from these sorted cells can be used to prepare libraries for RNAseq. Furthermore, we describe the main steps required to analyze our RNAseq results in order to show the reproducibility and relevance of the data. However, the primary aim of this protocol is not to describe a step-by-step method to analyze RNAseq results. Several papers have described how to align and map RNAseq results with different software [29, 30]. Other papers describe or compare several ways to perform a differential analysis to compare gene expression levels across different samples [31].

2 Materials

2.1 Cell Dissociation and FACS Sorting Components

1. Embryos from the zebrafish AB/TL control line 5 days post-fertilization (*see* **Note 1**).
2. Embryos from the transgenic zebrafish reporter line *Tg(mpeg1:Gal4-VP16)^{gl24};Tg(UAS-E1b:Kaede)^{s1999t}* 5 days post-fertilization [24].
3. Incubator (28.5 °C).
4. Egg water: "Instant Ocean" Sea Salts 60 μg/mL. 0.2 mM Phenylthiourea (PTU) (*see* **Note 2**).
5. Calcium-free Ringer solution: 5 mM HEPES (pH 7.2), 2.9 mM KCl, 116 mM NaCl.

6. Dissociation solution: Trypsin 0.25 % supplemented with 1 mM EDTA (Gibco®) (*see* **Note 3**).
7. Fetal calf serum (FCS) 100 % inactivated by heating for 30 min at 56 °C in a water bath with mixing.
8. $CaCl_2$ 0.8 M.
9. Dulbecco's Phosphate Buffered Saline (DPBS) 1×.
10. Resuspension solution: Leibovitz's L-15 medium + L-Glutamine without Phenol Red, FCS 10 %, 0.8 mM $CaCl_2$, penicillin 50 U/μL, streptomycin 0.05 mg/mL.
11. Sterile disposable 50 μm filters adaptable on Falcon tubes.
12. BD FACSAria™ III Cell Sorter (BD Biosciences, San Jose, CA, USA) with the BD FACSDiva software (version 6.1.3).
13. Cell collection solution: Leibovitz's L-15 medium + L-Glutamine without Phenol Red, FCS 10 %, zebrafish embryo extract 10 % (*see* **Note 4**), 0.8 mM $CaCl_2$, penicillin 50 U/μL, streptomycin 0.05 mg/mL.
14. 35 mm culture dishes.

2.2 RNA Extraction and Library Preparation Components

1. RNAqueous®-Micro Kit (Ambion®).
2. 100 % ethanol, ACS grade or better.
3. RNase-free low retention microcentrifuge tubes.
4. RNase-free filtered pipette tips.
5. Heating blocks at 75 and 37 °C.
6. Refrigerated microcentrifuge capable of at least 13,600 × g.
7. Agilent Bioanalyzer 2100, RNA 6000 Pico kit, DNA 1000 kit, and High sensitivity DNA kit (Agilent, Santa Clara).
8. Centrifugal Evaporator.
9. Clontech SMARTer™ Ultra Low RNA Kit for Illumina Sequencing (Clontech).
10. Agencourt AMPure XP (Beckman Coulter).
11. Magnetic rack for 1.5 mL microcentrifuge tubes.
12. PCR machine, and qPCR machine.
13. Covaris S220 Focused-ultrasonicator.
14. Illumina Truseq DNA Sample Preparation Kit v2 (Illumina Inc., San Diego, USA).
15. KAPA Library Quantification Kit (KAPA Biosystems).
16. Sequencing facility equipped with a HiSeq 2000 (Illumina Inc., San Diego, USA).
17. A set of software for RNAseq read analysis (*see* **Note 5**).

3 Methods

3.1 Cell Dissociation and FACS Sorting

1. Collect 150–200 5-day-old larvae grown in egg water in an incubator set up at 28.5 °C and transfer them into a 35 mm culture dish (*see* **Note 6**).
2. Rinse the larvae in 3 mL of calcium-free Ringer solution for 15 min.
3. Carefully remove as much as possible calcium-free Ringer solution and add 2–3 mL of dissociation solution pre-warmed at 28.5 °C.
4. Incubate for 90 min at 28.5 °C in an incubator. During incubation, grind up the larvae by pipetting up and down with a 1 mL tip for 10 min (*see* **Notes 7–9**).
5. Stop the reaction by adding $CaCl_2$ to a final concentration of 1 mM and fetal calf serum to 10 % (*see* **Note 10**).
6. Transfer cells into a microcentrifuge tube and centrifuge for 3 min at 800 × g (*see* **Note 11**).
7. Rinse the cells in 1 mL of DPBS and centrifuge again for 3 min at 800 × g.
8. Resuspend the cells in 1–1.5 mL of resuspension solution to obtain a concentration of 10^7 cells/mL (*see* **Note 12**).
9. Place a sterile disposable 50 μm filter on a 15 mL tube and load the cells onto it.
10. When all the liquid has passed through the filter, put the 15 mL cell-containing tube on ice and immediately proceed to cell sorting (*see* **Note 13**).
11. Subject the cell suspension to FACS for 20–30 min at 4 °C and collect the different cell fractions in microcentrifuge tubes containing 200 μL of cell collection solution (*see* **Notes 14** and **15**).
12. After cell sorting, the cells are kept on ice. Proceed to RNA extraction as soon as possible (*see* **Note 16**).

3.2 RNA Extraction, Library Preparation and RNA Sequencing

1. Pellet the cells by centrifugation at 13,000 × *g* for 4 min and proceed to RNA extraction.
2. For RNA extraction, use the RNAqueous®-Micro Kit and proceed according to the manufacturer's protocol.
3. At the end of the procedure, RNA extracted from nonfluorescent cells is resuspended in a final volume of 10 μL.
4. Before cDNA synthesis, remove DNA contaminants by DNase treatment using the DNase provided in the RNAqueous®-Micro Kit.

5. After DNase treatment, transfer RNA into low retention microcentrifuge tubes and store at −80 °C (*see* **Note 17**).
6. Before cDNA synthesis, measure RNA quantity and quality with an Agilent Bioanalyzer 2100 and RNA 6000 Pico kit (*see* **Note 18**).
7. If the RNA concentration is lower than 1,000 pg/μL, concentrate to a final volume of 1 μL using an Eppendorf Vacufuge set at 30 °C for approximately 10 min (*see* **Note 19**).
8. For each RNA sample, synthesize cDNA using the Clontech SMARTer Ultra Low RNA Kit for Illumina Sequencing according to the manufacturer's protocol (*see* **Note 20**).
9. Purify cDNA with AMPure XP beads according to the SMARTer Kit manual.
10. Verify quality and quantity of the cDNA syntheses by running 1 μL on an Agilent Bioanalyzer 2100 using the High sensitivity DNA kit (*see* **Note 21**).
11. Shear the amplified cDNA with a Covaris S220 system using microTUBE's and the settings recommended in the SMARTer Kit manual.
12. For each cDNA sample carry out library preparation with the Illumina TruSeq DNA Sample Preparation Kit v2 according to the manufacturer's protocol (*see* **Note 22**).
13. Verify library quality by running 1 μL of the libraries on an Agilent Bioanalyzer using the DNA 1000 kit (*see* **Note 23**).
14. Quantify the number of amplifiable molecules in the libraries using the KAPA Library Quantification Kit (KAPA Biosystems) (*see* **Note 24**).
15. Sequence the libraries on an Illumina HiSeq 2000 to obtain the desired number of paired end reads with a read length of 50 nucleotides (*see* **Notes 25** and **26**).

3.3 Data Analysis and Quality Assessment

1. Perform a quality trimming of the raw sequencing reads obtained from the CASAVA pipeline (Illumina Inc.) (*see* **Note 27**).
2. Align and map the uniquely mapping reads on the reference (*see* **Note 28**).
3. Check the reproducibility of biological replicates by calculating the square of the Pearson correlation coefficient (R^2) between all the counts or RPKM (read count per kilobase per million mapped reads) values from genes detected in both replicates (*see* **Notes 29** and **30**).
4. Perform differential expression analysis to detect significantly upregulated and downregulated genes in the fluorescence positive cells compared with the negative background (*see* **Note 31**).

4 Notes

1. Working with embryos and early larval stages does not require animal experimentation authorization. However, manipulation of larvae that have reached the free feeding stage and husbandry of adult fish require proper animal experimentation authorization according to standard regulations in each country.
2. Pigmentation of the larvae does not interfere with fluorescent-based cell sorting. If required, 0.2 mM Phenylthiourea (PTU) can be added to the egg water in order to prevent melanization and allow screening of transgenic larvae. However, one should notice that addition of PTU in egg water can interfere with biological functions [32, 33].
3. Trypsin reagents form different suppliers were tested for dissociation of zebrafish larvae. We obtained the best dissociation efficiency and cell survival with Trypsin 0.25 % supplemented with 1 mM EDTA from Gibco®.
4. Stock of zebrafish embryo extract obtained from 200 AB/TL larvae was prepared according to the protocol from the zebrafish book [34].
5. Various software programs, charged or free, are available. We have used Illumina HCS version 1.15.1 for image analysis and base calling, CLCbio Assembly Cell v4.0.6 for quality trimming of sequence reads and mapping of filtered reads to Ensembl transcripts, and the DEseq package (version 1.8.3; [35]) available in Bioconductor (version 2.10) for analysis of differential gene expression between fluorescence positive and negative cells. Computer analyses will not be explained in detail in this protocol because it requires a complete bioinformatic protocol to explain each step of the analysis.
6. The number of larvae one needs to collect depends on the number of fluorescent cells expected to be collected after FACS sorting. With the *Tg(mpeg1:Gal4-VP16);Tg(UAS-E1b:Kaede)* line, 0.1 % of the cells are fluorescence positive and using 200 fish allows to obtain an average of 6,000 cells.
7. Incubation time depends on the number of larvae and also their age. An incubation time of 90 min at 28.5 °C works well with 150–200 larvae collected 5 or 6 days post-fertilization. If more than 200 embryos need to be dissociated simultaneously, sample should be separated into two different culture dishes filled up with 2 mL Trypsin 0.25 % each.
8. A dissociation time too short leads to a decrease in the number of cells obtained after FACS sorting. A dissociation time too long is harmful for the cells. In order to be sure to have a perfectly homogenous mixture of single cells, it is particularly

important to disintegrate cell aggregates that may form during incubation at 28.5 °C.

9. The influence of trypsin treatment on the transcriptome is difficult to assess. In order to minimize artefactual signal, dissociation of all the samples should be performed under the same conditions (number of larvae, dissociation time, trypsin reagent batch …).
10. Stop the reaction when the solution appears as fluid as water. One can also assess the dissociation by observing a mainly single cell suspension under a microscope.
11. One should observe a pellet of cells with a silvery color. If there are too many cells, some of them will remain in the supernatant. In this case, centrifuge the supernatant again and collect the remaining cells.
12. A cell count has been performed during the first experiments. Based on these results we found that resuspending dissociated cells from 200 5-day-old larvae in 1 mL allows obtaining the desired concentration of 10^7 cells/mL.
13. This step allows collecting only dissociated single cells. If the cells are not well dissociated, the filter could become blocked. If this happens, shaking the filter carefully may help the liquid pass through the filter. If this does not work, load the remaining liquid on a new filter.
14. We have used a FACSAria III (BD Biosciences) with the BD FACSDiva software (version 6.1.3). To sort Kaede green positive cells a Coherent Sapphire solid-state 488 nm laser with 15.4 mW power was used. Laser settings applied were 505 LP, 530/30 BP. In order to set up sorting gates, we have previously sorted single cell suspensions from 5 dpf AB/TL control larvae obtained with the same protocol. Gates are set up in order to exclude all autofluorescent cells. Sorting of additional cell suspensions showed that no more than 10–30 false positive cells are sorted per 150 AB/TL larvae with these predefined gates. In contrast, we were routinely able to obtain more than 6,000 positive cells from 150 *Tg(mpeg1:Gal4-VP16);Tg(UAS-E1b:Kaede) larvae*.
15. As many cells as possible are collected for the fluorescent populations and a maximum of 500,000 cells for the negative fraction. Cell collection solution allows cell survival [36]. However, FACS should not be performed for more than 30 min per sample in order to avoid damaging the cells.
16. The purity of sorted cells can be assessed by fluorescence microscopy and by performing cytospin preparations followed by Giemsa staining of cells. For cytospin analysis, cells should be maintained in culture for a few hours after FACS because

cells adopt a round-shaped morphology after trypsin treatment making it impossible to differentiate the immune cell types.

17. Transferring RNA into low retention microcentrifuge tubes is useful to avoid RNA loss by binding to plastic tubes. This is particularly relevant for low concentrated RNA kept for several weeks at −80 °C.
18. RNA quality is reflected by the RNA Integrity Number (RIN) provided at the end of the run on the bioanalyzer. RNA samples with a RIN comprised between 7 and 10 are usually requested by sequencing platforms. Low RNA concentrations (<200 pg/μL) can sometimes not be estimated with the bioanalyzer and therefore no RIN is associated with these samples. Nevertheless, if the peaks corresponding to ribosomal RNA are detected, these samples can be used to obtain a good quality library as shown below (Fig. 2).
19. Sometimes a concentration step may be harmful for RNA. If a degradation of RNA is noticed during this step, adding RNase inhibitors has been described as an efficient way to reduce RNA degradation [37].
20. A protocol has been recently described to perform cDNA synthesis with the Clontech SMARTer Ultra Low RNA Kit for Illumina Sequencing followed by library sequencing with one cell only [6, 8]. With this protocol, a unique cell is lysed directly in cDNA synthesis buffer. It seems likely than this protocol could be further adapted for a small number of cells obtained by FACS or microdissection.
21. A good amplification product should produce an electropherogram with an increase in fluorescence intensity comprised between 400 and 10,000 bp, with a maximum intensity at 1,000–2,000 bp (Fig. 2a, b). A signal detected between 0 and 250 bp indicates the presence of poly-A stretches. If poly-A stretches are present but in minority compared to the amount of amplified cDNA, the sample can be used for library synthesis. Absence of fluorescence between 400 and 10,000 bp indicates that synthesis of cDNA has failed, whereas the presence of several discrete peaks suggests a contamination of the sample. If the synthesis of cDNA fails, carefully check that no degradation of RNA has occurred when concentrating the RNA. For more information one can refer to Clontech SMARTer™ Ultra Low RNA Kit for Illumina Sequencing handbook.
22. Compared to the manufacturer's protocol, only two modifications have been made. In the adapter ligation step the adapters were diluted 20-fold, and in the amplification step fifteen cycles were used instead of ten. These modifications have been found to increase efficiency of cDNA synthesis.

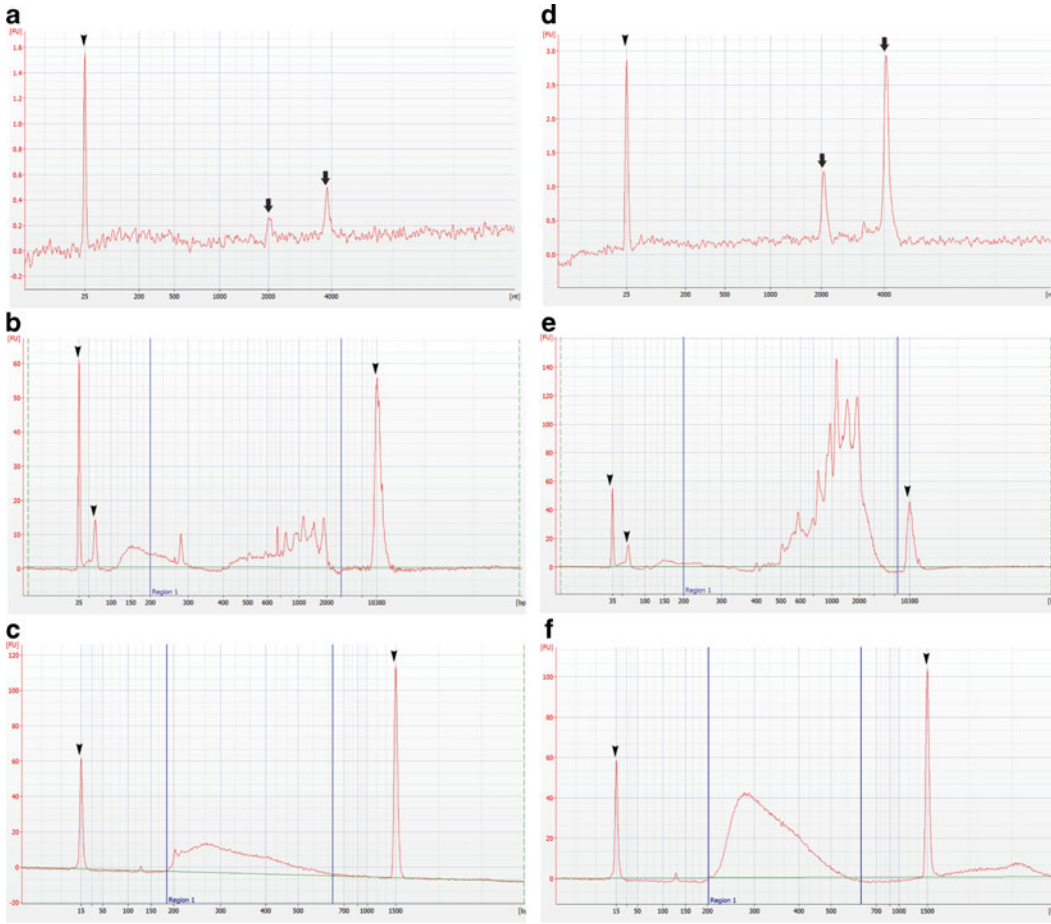

Fig. 2 Electropherograms of extracted RNA, amplified cDNA, and final libraries. Each step of the library preparation is checked by running 1 μL on an Agilent Bioanalyzer 2100. *X*-axis represents RNA or DNA fragment size in nucleotides (nt) or base pair (bp), respectively. *Y*-axis represents intensity of fluorescence detected (FU). Representative results are shown for two RNA samples obtained from *Tg(mpeg1:Gal4-VP16);Tg(UAS-E1b:Kaede)* positive cell fractions. (**a**–**c**) Results obtained during preparation of sample 1. (**d**–**f**) Results obtained during preparation of sample 2. (**a**, **d**) Electropherograms of RNA obtained after extraction from FACS sorted cells with the RNAqueous®-Micro Kit after DNase treatment. (**b**, **e**) Electropherograms of amplified cDNA obtained with Clontech SMARTer™ Ultra Low RNA Kit for Illumina Sequencing after 15 amplification cycles. (**c**, **f**) Electropherograms of libraries obtained after library preparation from the corresponding cDNA using the Illumina TruSeq DNA Sample Preparation Kit v2. RNA concentrations for samples 1 and 2 are 47 and 144 pg/μL, respectively. In both cases the presence of ribosomal RNA peaks in the electropherograms (*arrows* in **a**, **d**) indicated the integrity of the RNA sample, but a RIN value failed to be calculated. These low amounts of RNA from fluorescence positive cells gave rise to cDNA preparations with a broad size distribution (areas in between the *blue lines* in **b** and **e**) and Illumina TruSeq libraries (areas in between the *blue lines* in **c** and **f**) that were successfully applied to RNAseq with a HiSeq 2000 system. With these libraries, we obtained 10 million reads per sample, which was sufficient to detect an average of 11,000 expressed genes (with RPKM values $\geq$0.5) in fluorescence positive cells on a total of 27,882 genes present in our reference. Peaks indicated by *arrowheads* at the start (**a**–**f**) and end (**b**, **c**, **e**, **f**) of the electropherograms are size markers used for calibration of the Bioanalyzer

23. A good library preparation gives a fluorescence absorption distributed between 200 and 600 bp with a maximum intensity around 200–300 bp. Absence of fluorescence between 200 and 600 bp indicates a failure in library preparation whereas a peak of fluorescence shifted either to the low or high DNA size reveals too much shearing or not enough shearing of cDNA, respectively.
24. Library quantification is primordial to load the optimal amount of DNA into the sequencing flow cell and thus achieving an optimal sequencing result.
25. Sequencing 10 million paired end reads per sample was sufficient to detect differential expression of macrophage markers (e.g., *mpeg1*, *mhc2dab*, *mfap4*, *csf1ra*, *marco*, *irf8*) between fluorescence positive and fluorescence negative cells. However, more reads (100 million) are recommended to detect rare transcripts and to analyze alternative transcription start sites, splicing or polyadenylation.
26. Library preparation and sequencing (**steps 6–15**) are often carried out by sequencing facilities. However, library preparation from RNA extracted from FACS sorted cells differs from regular protocols and should be carefully discussed with the sequencing facility.
27. For this analysis we used the quality trim option available in CLC Assembly Cell v4.0.6 beta (CLC bio, Denmark) with standard settings. Alternatively, one can use the Filter FASTQ option freely available in the Galaxy pipeline [38].
28. To align and map reads we used CLC Assembly Cell 4.0.6 beta (CLC bio, Denmark). Reads were mapped to Ensembl transcripts (Zv9_63) using the clc_ref_assemble_short module. Transcripts were then accumulated to their corresponding Ensembl gene using the assembly_table module. Reads belonging to the same gene were finally summed together using a custom perl script. Alternatively one can use other aligners such as Bowtie [39] or TopHat [40] and the genomic databases available on Ensembl.
29. The square of the Pearson coefficient (R^2) reflecting the linear correlation between the RPKM of two samples is used to assess the reproducibility of these two independent experiments. Our results gave a R^2 higher than 0.9 for the reads obtained from the fluorescence negative cells. For the fluorescence positive cells, of which the number of cells obtained after FACS is low, R^2 values between 0.867 and 0.964 were obtained (Fig. 1). Thus, our results show good correlation, but should be reproduced at least in three replicates in order to obtain significant results with the differential expression analysis.

30. We noticed that incorporating nonunique mapped reads in our analysis dramatically decreased the Pearson correlation coefficient between our replicates. Thus, we strongly recommend working with unique mapped reads.
31. To perform statistical analyses we used the R package DESeq [35] and selected genes upregulated or downregulated more than twofold with an adjusted *p*-value smaller than 0.1. Alternatively, differential analyses can be performed with other R packages such as EdgeR [41] or Bayseq [42].

Acknowledgments

This work was supported by the Marie Curie Initial Training Network FishForPharma (PITN-GA-2011-289209) and the project ZF-HEALTH (HEALTH-F4-2010-242048) funded by European Commission 7th Framework Programme, and by the SmartMix programme of the Netherlands Ministry of Economic Affairs and the Ministry of Education, Culture, and Science. Additionally, Z.K. was supported by the Higher Education Commission of Pakistan, and A.Z. was supported by a Horizon grant of the Netherlands Genomics Initiative.

References

1. Westermann AJ, Gorski SA, Vogel J (2012) Dual RNA-seq of pathogen and host. Nat Rev Microbiol 10(9):618–630
2. Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10(1):57–63
3. Fullwood MJ, Wei CL, Liu ET et al (2009) Next-generation DNA sequencing of paired-end tags (PET) for transcriptome and genome analyses. Genome Res 19(4):521–532
4. Trapnell C, Williams BA, Pertea G et al (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 28(5):511–515
5. Mortazavi A, Williams BA, McCue K et al (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5(7):621–628
6. Cann GM, Gulzar ZG, Cooper S et al (2012) mRNA-Seq of single prostate cancer circulating tumor cells reveals recapitulation of gene expression and pathways found in prostate cancer. PLoS One 7(11):e49144
7. Lao KQ, Tang F, Barbacioru C et al (2009) mRNA-sequencing whole transcriptome analysis of a single cell on the SOLiD system. J Biomol Tech 20(5):266–271
8. Ramskold D, Luo S, Wang YC et al (2012) Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat Biotechnol 30(8):777–782
9. Tang F, Barbacioru C, Nordman E et al (2010) RNA-Seq analysis to capture the transcriptome landscape of a single cell. Nat Protoc 5(3): 516–535
10. Meijer AH, Spaink HP (2011) Host-pathogen interactions made transparent with the zebrafish model. Curr Drug Targets 12(7):1000–1017
11. Ramakrishnan L (2012) Revisiting the role of the granuloma in tuberculosis. Nat Rev Immunol 12(5):352–366
12. Renshaw SA, Trede NS (2012) A model 450 million years in the making: zebrafish and vertebrate immunity. Dis Model Mech 5(1): 38–47
13. van der Vaart M, Spaink HP, Meijer AH (2012) Pathogen recognition and activation of the innate immune response in zebrafish. Adv Hematol 2012:159807
14. Davis JM, Clay H, Lewis JL et al (2002) Real-time visualization of mycobacterium-

macrophage interactions leading to initiation of granuloma formation in zebrafish embryos. Immunity 17(6):693–702

15. Levraud JP, Disson O, Kissa K et al (2009) Real-time observation of Listeria monocytogenes-phagocyte interactions in living zebrafish larvae. Infect Immun 77(9):3651–3660
16. Prajsnar TK, Hamilton R, Garcia-Lara J et al (2012) A privileged intraphagocyte niche is responsible for disseminated infection of Staphylococcus aureus in a zebrafish model. Cell Microbiol 14(10):1600–1619
17. van der Sar AM, Spaink HP, Zakrzewska A et al (2009) Specificity of the zebrafish host transcriptome response to acute and chronic mycobacterial infection and the role of innate and adaptive immune components. Mol Immunol 46(11–12):2317–2332
18. Vergunst AC, Meijer AH, Renshaw SA et al (2010) Burkholderia cenocepacia creates an intramacrophage replication niche in zebrafish embryos, followed by bacterial dissemination and establishment of systemic infection. Infect Immun 78(4):1495–1508
19. Encinas P, Rodriguez-Milla MA, Novoa B et al (2010) Zebrafish fin immune responses during high mortality infections with viral haemorrhagic septicemia rhabdovirus. A proteomic and transcriptomic approach. BMC Genomics 11:518
20. Hegedus Z, Zakrzewska A, Agoston VC et al (2009) Deep sequencing of the zebrafish transcriptome response to mycobacterium infection. Mol Immunol 46(15):2918–2930
21. Ordas A, Hegedus Z, Henkel CV et al (2011) Deep sequencing of the innate immune transcriptomic response of zebrafish embryos to Salmonella infection. Fish Shellfish Immunol 31(5):716–724
22. Stockhammer OW, Rauwerda H, Wittink FR et al (2010) Transcriptome analysis of Traf6 function in the innate immune response of zebrafish embryos. Mol Immunol 48(1–3): 179–190
23. Yang D, Liu Q, Yang M et al (2012) RNA-seq liver transcriptome analysis reveals an activated MHC-I pathway and an inhibited MHC-II pathway at the early stage of vaccine immunization in zebrafish. BMC Genomics 13:319
24. Ellett F, Pase L, Hayman JW et al (2011) mpeg1 promoter transgenes direct macrophage-lineage expression in zebrafish. Blood 117(4): e49–e56
25. Hall C, Flores MV, Chien A et al (2009) Transgenic zebrafish reporter lines reveal conserved Toll-like receptor signaling potential in embryonic myeloid leukocytes and adult immune cell lineages. J Leukoc Biol 85(5): 751–765
26. Langenau DM, Ferrando AA, Traver D et al (2004) In vivo tracking of T cell development, ablation, and engraftment in transgenic zebrafish. Proc Natl Acad Sci U S A 101(19): 7369–7374
27. Renshaw SA, Loynes CA, Trushell DM et al (2006) A transgenic zebrafish model of neutrophilic inflammation. Blood 108(13): 3976–3978
28. Wittamer V, Bertrand JY, Gutschow PW et al (2011) Characterization of the mononuclear phagocyte system in zebrafish. Blood 117(26):7126–7135
29. Garber M, Grabherr MG, Guttman M et al (2011) Computational methods for transcriptome annotation and quantification using RNA-seq. Nat Methods 8(6):469–477
30. Trapnell C, Roberts A, Goff L et al (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7(3):562–578
31. Dillies MA, Rau A, Aubert J, Hennequet-Antier C, Jeanmougin M, Servant N, Keime C, Marot G, Castel D, Estelle J, Guernec G, Jagla B, Jouneau L, Laloë D, Le Gall C, Schaëffer B, Le Crom S, Guedj M, Jaffrézic F (2012) French StatOmique Consortium. A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. Brief Bioinform. 14(6): 671–83
32. Elsalini OA, Rohr KB (2003) Phenylthiourea disrupts thyroid function in developing zebrafish. Dev Genes Evol 212(12):593–598
33. Li Z, Ptak D, Zhang L et al (2012) Phenylthiourea specifically reduces zebrafish eye size. PLoS One 7(6):e40132
34. Westerfield M (1993) The zebrafish book: a guide for the laboratory use of zebrafish (Brachydanio rerio). M. Westerfield, Eugene, OR
35. Anders S, Huber W (2010) Differential expression analysis for sequence count data. Genome Biol 11(10):R106
36. Covassin L, Amigo JD, Suzuki K et al (2006) Global analysis of hematopoietic and vascular endothelial gene expression by tissue specific microarray profiling in zebrafish. Dev Biol 299(2):551–562
37. Cummings M, McGinley CV, Wilkinson N et al (2011) A robust RNA integrity-preserving staining protocol for laser capture microdissection of endometrial cancer tissue. Anal Biochem 416(1):123–125

38. Blankenberg D, Gordon A, Von Kuster G et al (2010) Manipulation of FASTQ data with Galaxy. Bioinformatics 26(14):1783–1785
39. Langmead B, Trapnell C, Pop M et al (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10(3):R25
40. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25(9):1105–1111
41. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26(1): 139–140
42. Hardcastle TJ, Kelly KA (2010) baySeq: empirical Bayesian methods for identifying differential expression in sequence count data. BMC Bioinformatics 11:422