

# NIH Public Access

**Author Manuscript**

Curr Protoc Hum Genet. Author manuscript; available in PMC 2013 June 05.

Published in final edited form as:

Curr Protoc Hum Genet. 2012 April ; 0 11: Unit–11.1210. doi:10.1002/0471142905.hg1112s73.

## **High-throughput multiplex sequencing of miRNA**

**Francois Vigneault**1,2,3,\* , **Dmitry Ter-Ovanesyan**2,4,\* , **Shahar Alon**5,6,\* , **Seda Eminaga**1, Danos C. Christodoulou<sup>1</sup>, J. G. Seidman<sup>1</sup>, Eli Eisenberg<sup>5,6</sup>, and George M Church<sup>1,2</sup> <sup>1</sup>Department of Genetics, Harvard Medical School, Boston, MA 02115, USA

<sup>2</sup>Wyss Institute for Biologically Inspired Engineering, Boston, MA 02115, USA

<sup>3</sup>Ragon Institute of MGH, MIT and Harvard, Massachusetts, MA 02129, USA

<sup>4</sup>Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02139, USA

<sup>5</sup>Department of Neurobiology, George S. Wise Faculty of Life Sciences, Tel-Aviv University, Israel

<sup>6</sup>Raymond and Beverly Sackler School of Physics and Astronomy, Tel-Aviv University, Israel

#### **Abstract**

Next-Generation Sequencing offers many advantages over other methods of microRNA (miRNA) expression profiling, such as sample throughput and the capability to discover novel miRNAs. As the sequencing depth of current sequencing platforms exceeds what is necessary to quantify miRNAs, multiplexing several samples in one sequencing run offers a significant cost advantage. Although previous studies have achieved this goal by adding barcodes to miRNA libraries at the ligation step, this was recently shown to introduce significant bias into the miRNA expression data. This bias can be avoided, however, by barcoding the miRNA libraries at the PCR step instead. Here, we describe a user-friendly PCR bar-coding method of preparing multiplexed microRNA libraries for Illumina-based sequencing. The method also prevents the production of adapter dimers and can be completed in one day.

#### **Keywords**

miRNA; Illumina; Sequencing; library; multiplex; bar code

### **INTRODUCTION**

MicroRNAs (miRNAs) are approximately 21–23 nucleotide RNAs that control the expression of most genes. They are involved in many normal cellular processes and their dysregulation has been implicated in many diseases (Kloosterman & Plasterk, 2006). Thus, profiling miRNA expression is important for understanding biological processes, as well as for the development of new diagnostic signatures and therapeutic targets. Next Generations Sequencing (NGS) is currently the most comprehensive method for profiling miRNA expression. In addition to avoiding issues with background and non-specific hybridization

#### **Competing financial interests**

Correspondence should be addressed to: Francois Vigneault, fvigno@hms.harvard.edu.

<sup>\*</sup>These authors contributed equally to this work

**Author Contributions**

F.V. designed and produced the microRNA library protocol and wrote the manuscript, D.T. fine-tuned the protocol and wrote the manuscript, S.A. and D.C. provided bioinformatics analysis, S.E. provided samples and assistance in developing the protocol, J.G.S., E.E. and G.M.C. supervised all parts of the work.

The authors declare no competing financial interests.

found in alternative methods such as miRNA microarrays and qRT-PCR, NGS also allows for the identification of new miRNAs or other small regulatory RNAs (Creighton, Reid, & Gunaratne, 2009).

The combining of multiple biological samples in one sequencing reaction, called multiplexing, is a popular method of reducing cost for NGS. This can be done by attaching a specific sequence tag (bar-code) to each sample before combining them in one sequencing run. In this protocol, we describe a quick and cost-effective method to make bar-coded miRNA libraries for NGS, which can be completed in a single day.

#### **BASIC PROTOCOL**

#### **MULTIPLEX microRNA LIBRARY CONSTRUCTION FOR ILLUMINA SEQUENCING**

This procedure describes a method for constructing multiplexed miRNA libraries. A miRNA library is made (Figure 1) from each RNA sample by 3' adapter ligation, 5' RT primer annealing, 5' adapter ligation, reverse transcription, and PCR amplification. Although the forward PCR primer is the same, a different reverse PCR primer with a unique barcode is used for each RNA sample. The different libraries can then be pooled into a single sequencing reaction at the end of the library construction. The following instructions are for the preparation of one sample, so users must scale-up according to the specific number of samples they are preparing. All incubations are conducted in a thermal cycler.

#### **MATERIALS**

RNase Zap (Ambion, AM9780) Nuclease-free water (Ambion, AM9937) 10x T4 RNA Ligase 2tr (Enzymatics, L607) 10x T4 RNA Ligase 2tr Buffer (Enzymatics, L607) 100% DMSO (Sigma, D9170) RNase Inhibitor (Enzymatics, Y924L) ATP (Enzymatics, N207-10-L) T4 RNA Ligase 1 (Enzymatics, L605L) dNTPs (Enzymatics, N205L) Superscript III First-Strand Synthesis System (Invitrogen, 180080-051) Phusion High-Fidelity DNA Polymerase (NEB, M0530S) AgencourtAMPure XP 5 mL Kit (Beckman Coulter Genomics, A63880) E-Gel EX Gel, 2% (Invitrogen, G4020-02) Ethanol 25 bp Ladder (Invitrogen, 10597-011) 100 bp Ladder (Invitrogen, 15628-019) MinElute Reaction Cleanup Kit (Qiagen, 28204) Agilent High Sensitivity DNA Kit (Agilent, 5067-4626) Thermal Cycler (for all incubations) E-Gel I-Base Power System (Invitrogen, G6400)

E-Gel Safe Imager Real-Time Transilluminator (Invitrogen, G6500) Dynamag-2 Magnet (Invitrogen, 123-21D) Recommended: Iceless Cold Pack (Eppendorf 022510509) Recommended: Agilent 2100 Bioanalyzer Optional: Nanodrop Spectrophotometer 2000

#### **PROCEDURE**

**Ligation of 3' adenylated adapter—**Make sure to clean surfaces and instruments with RNase Zap and maintain RNase-free conditions throughout the protocol. While as few as 100 ng of total RNA is sufficient, we recommend starting with at least 1 µg of total RNA (one can also use the equivalent fraction of enriched for small RNAs if desired). We recommend verifying RNA quality using the Agilent Bioanalyzer RNA nano or pico chip and using samples of RIN value of 7 or above.

- 1| Dilute the starting RNA to 200 ng/ $\mu$ L in nuclease-free dH<sub>2</sub>O, if possible.
- **2|** Set up ligation reaction in a PCR tube.



**3|** Denature at 90°C for 30 sec, then 4°C for at least 30 sec.

**4|** Add the following directly to the ligation reactions on ice:

<b>Component</b>	Volume $(\mu\mathbf{l})$	<b>Final Concentration</b>
RNase Inhibitor $(40 \text{ U/}\mu\text{I})$	0.5	2 U/ul
T4 RNA Ligase $2tr(200 U/\mu l)$	1.5	30 U/ul

**<sup>5|</sup>** Incubate at 22°C for 1 h.

We recommend using Enzymatics buffer as its composition gave us significantly higher yield than other commercially available T4 RNA ligase 2 truncated buffers.

#### **Annealing of RT primer—**

**6|** Add the following directly to each reaction on ice:



The final amount of RT primer must be at equimolar ratio (10 pmoles) with the starting amount of 3'rApp-adapter (10 pmoles) for each sample.

**7|** Incubate at 90°C for 30 sec, then 65°C for 5 min, then 4°C for at least 30 sec.

#### **Ligation of 5' RNA adapter—**

**8|** Prepare the 5' RNA adapter by incubating ~5µl at 70°C for 2 min, then 4°C for at least 30 sec.

An excess of volume is prepared to account for evaporation and facilitate the pipetting of the proper volume at the next step.

**9|** Spin down the ligation mixture and add the following reagents directly to it:

Component	Volume (µI)	<b>Final Concentration</b>
10 mM ATP	1.5	$1 \mu M$
10 µM 5' RNA Adapter		10 pmoles total
T4 RNA Ligase $1(20 U/\mu l)$	1.5	$2 U/ \mu l$

**<sup>10|</sup>** Incubate at 20°C for 1 h.

**Reverse transcription of captured MicroRNAs—**The previous steps result in a reaction volume of 15 µl. Only 5µl is used in the subsequent RT-PCR step, and so the remaining can be stored (−80°C) as a backup (highly recommended) for two more runs. However, the rest of the protocol below can be scaled up 3 times and the full 15µl may be processed at once if you need to achieve higher yield (for example, when starting with lower amounts of RNA).

**11** Prepare the following reaction in a PCR tube:



**12|** Incubate at 48°C for 30 min.

As the RT primer was annealed earlier, do not denature or conduct an annealing cycle at this stage but go directly to the reverse transcriptase incubation temperature (48°C, as shown above).

#### **Limited PCR Amplification—**

**13|** Prepare the PCR reaction in a PCR tube:



BcmiRNA\_PCR2\_BC\* refers to the bar-coded primer, where for each unique starting RNA sample a unique bar-code primer needs to be used (see Table 1). To limit bar-code / samples aerosol contamination, it is recommended to only open and close one tube of primer at a time.

- **14|** Cycle the PCR reaction as follow in a thermal cycler:
	- 1- 98°C for 30 sec
	- 2- 98°C for 10 seconds
	- 3- 60°C for 20 seconds
	- 4- 72°C for 20 seconds go to step 2, 11 more time
	- 5- 72°C for 5 min
	- 6- 4°C pause

The number of cycles can be varied according to the amount of microRNA present in the starting sample. In our hands, a total of 12 cycles generally results in the best yield while limiting unnecessary cycling. We recommend not exceeding 15 cycles as this will increase non-specific background amplification and reduce optimal yield of the desired products. Instead, additional starting RNA should be prepared in parallel and combined at the final stage to increase yield.

#### **PCR Clean-Up with AMPure XP beads—**

- **15|** Transfer PCR reactions to a new 1.5 mL tube.
- **16|** Vigorously mix the AMPure XP beads and then add 90 µl of beads to each 50 µl PCR reaction. Pipet the beads slowly.
- **17|** Vortex for 30 seconds, and then incubate on bench for 5 min.
- **18|** Quick spin, and place on the magnetic rack for 5 min.
- **19|** With the tubes still on the magnet, aspirate and discard the liquid from the reaction.
- **20|** With the tubes still on the magnet, add 400 µL 70% EtOH to the beads and leave for 30 sec. Then discard the 70% EtOH.
- **21|** Repeat the previous step for a second wash.
- **22|** Quick spin on a microfuge to collect last traces of EtOh. Put tubes back on magnet and remove any last drops of EtOH at the bottom of the tube.

- **23|** Leave the tube open to air dry for 2 min.
- **24|** Remove the tube from the magnet and add 45 µL of nuclease-free water.
- **25|** Vortex for 30 sec.
- **26|** Place the tubes on the magnet and leave for 1 min.
- **27|** With the tubes on the magnet, transfer 42 µL to new 1.5 mL tubes.

#### **Gel extraction of microRNA library—**

- **28|** Prepare a 2% Agarose Gel EX following the manufacturer's protocol.
- **29|** Dilute the 25 bp and 100 bp ladders 1:20 in water and load 20 µL of each.
- **30|** Split each microRNA library prepared above across 2 lanes by loading 20 ul per well.
- **31|** Run the gel for 14 minutes on the Invitrogen I-Base using the 2% E-Gel settings.
- **32|** When the run is complete, take a picture of the gel.

The migration patterns of DNA on E-gels are affected by the total amount and salts present in the loaded sample, and sometime one may observe a shift in migration of the expected product in relation to the ladder.

- **33|** Pry open the E-Gel by cracking open each side.
- **34|** Using a clean razor blade for each sample, cut between 125 and 175 bp to capture the two dominant miRNA bands.
- **35|** Gel extract using the Mini Elute Qiagen Gel-Extraction Kit following the manufacturer's protocol, conducting the final elution in 15 ul of dH2O.

Melt the gel bands at  $37^{\circ}$ C instead of the recommended 55  $^{\circ}$ C. The MinElute columns have a tendency to trap residual EtOH from the wash steps. To avoid this issue, dry spin the column for 1 minute at maximum speed and then turn the column 180 degrees and repeat the spin for another 1 minute. Then transfer the column to a recovery tube and leave the column open for 3 min to air dry prior to adding the elution buffer.

#### **Library QC and Mixing—**

**36|** The library can now be mixed at equimolar concentration, prior to submission for sequencing. We strongly recommend analyzing the quality and concentration of each final library using the Agilent Bioanalyzer DNA high sensitivity chip in order to combine the different libraries at equimolar ratios into a single multiplexed library. Although less accurate, a Nanodrop spectrophotometer (APPENDIX 3D) would also work to a decent degree for this step if an Agilent Bioanalyzer is inaccessible. For high throughput project with high amount of samples, the Bioanalyzer can be used to combine the libraries prior gel extraction.

#### **COMMENTARY**

#### **BACKGROUND INFORMATION**

Most commercially-available kits for NGS have followed the same general method (Thomas & Ansel, 2010) of preparing miRNA libraries, similar to that used to make a miRNA library for shotgun sequencing (Lau, Lim, Weinstein, & Bartel, 2001). First, a 3' adenylated DNA adapter is ligated to the total pool of miRNAs in the absence of ATP. Using an adenylated

adapter prevents the self-circularization of miRNAs. Then, a 5' RNA adapter is ligated to the miRNAs in the presence of ATP. Reverse transcription (RT) is performed with a primer complimentary to the 3' adapter. Finally, the captured miRNAs are PCR amplified with primers complementary to the 5' and 3' adapters. The PCR product is run out by gel electrophoresis and the appropriate band can be excised. One issue with this approach is that excess of 5' adapter can ligates directly to the excess of 3' adapter, and this adapter dimer gets PCR amplified to be the predominant product in the library.

We have previously reported that one efficient way of removing this adapter dimer was to conduct two denaturing PAGE extractions (Alon et al., 2011), but this is quite labor intensive and time consuming. Other methods have been developed to reduce the formation of this adapter dimers but with limited success (Kawano et al., 2010). Similarly to others (NEB product E6120 and Bioscientific product 5132-02), we have noted that pre-annealing the RT primer in between the 3' adapter ligation and 5' adapter ligation steps, as opposed to doing so at the RT step, works efficiently to reduce adapter dimer. This works by making the free 3' adapter double stranded, thus preventing it from ligating to the single-stranded 5' adapter.

Our protocol includes multiplexing of several miRNA samples in one sequencing reaction with bar-coded PCR primers. Considering the high sequencing capacity of one lane of a flow cell on an Illumina machine, multiplexing can significantly reduce the cost of miRNA profiling. Although some groups have previously used barcoded 3' or 5' oligonucleotide adapters to multiplex their miRNA libraries (Tarasov et al., 2007; Uziel et al., 2009; Vigneault et al., 2008; J. Y. Zhu et al., 2009; Q.-H. Zhu et al., 2008), we and others have shown that this leads to significant bias in the miRNA expression profile (Alon et al., 2011; Hafner et al., 2011). Instead, introducing bar-codes during the PCR amplification step, as described in this protocol, can eliminate this bias (Alon et al., 2011).

#### **CRITICAL PARAMETERS AND TROUBLESHOOTING**

Although we have not seen an adapter dimer band (114 bp) on an agarose gel using this protocol, if a bright adapter dimer band is present, we recommend performing one or to two denaturing PAGE extractions if required, as describe previously (Alon et al., 2011). When visible, most of the adapter dimer can be extracted under native condition, but a fraction can also anneal to the full-length library fragments (creating a mistmatch at the miRNA location), which is the carried over during extraction of the desired product. This is because the microRNA represent only a small portion  $\left(\sim 22bp\right)$  of the final library size  $\left(\sim 135bp\right)$ , and the majority of the remaining sequence is common to any other library fragments as well as the adapter dimer. To bypass this issue, the sample can be run on a denaturing gel in order to efficiently remove such undesirable bands. If the adapter dimer is still visible after one extraction, a second usually remove any residual trace.

This library design allows for sequencing the miRNA on both the Illumina HiSeq and GaIIX. Our preferred approach is to conduct a single pass sequencing of  $\sim$ 75 bp where the miRNA is first sequenced, followed by a short portion of the 3' adapter and then the barcode (as depicted in Figure 1). This is done using the standard Illumina sequencing primer for genomic libraries (or TruSeq primer), making it convenient to share the same flow cell with other type of libraries. Optionally, a custom indexing primer can also be used (see Table 1), but this is usually a more expensive option as one as to cycle the full flow cell to do so, which is rarely a requirement for most other customer sharing the run.

#### **ANTICIPATED RESULTS**

Our protocol describes the preparation of multiplexed miRNA libraries for NGS. Following library preparation, bar-coded samples are loaded on a 2% agarose gel (Figure 2), the miRNA fraction can then be separated from other RNAs by cutting the gel between 125 bp and 175 bp. There should be little or no amplified adapter-dimer product (114 bp) visible on the gel (Figure 3). Following gel extraction, the different bar-coded samples are mixed in equal proportion and sequenced in one reaction on an Illumina machine (GAII, HiSeq2000). A standard Illumina sequencing primer can be used to sequence both the miRNA and the bar-code in one sequencing pass (we recommend using 75 bp sequencing pass) or the bar code can be sequenced separately with a different custom indexing primer.

#### **TIME CONSIDERATIONS**

This protocol can be completed in one day.

Steps 1–5: 1 Hour 30 Minutes

Steps 6–7: 10 Minutes

Steps 8–10: 1 Hour 15 Minutes

Steps 11–12: 45 Minutes

Steps 15–27: 20 Minutes

Steps 28–35: 1 Hour

#### **Acknowledgments**

This work was supported by the Center for Excellence in Genome Sciences grant from the National Human Genome Research Institute. F.V. is supported by a Canadian Institutes of Health Research and Ragon Institute Fellowship. J.G.S is supported by grants from the NIH, NHLBI, the SysCODE Consortium (NIH) and the FondationLeducq. D.T. is supported by the Mazur Fellowship and the NIH Training Grant "Joint Program in Cells, Molecules, and Organisms." This work was partially supported by a grant from the United States-Israel Binational Science Foundation (grant no. 2009290), Jerusalem, Israel.

#### **Literature Cited**

- Alon S, Vigneault F, Eminaga S, Christodoulou DC, Seidman JG, Church GM, Eisenberg E. Barcoding bias in high-throughput multiplex sequencing of miRNA. Genome research. 2011; 21(9): 1506–1511. [PubMed: 21750102]
- Creighton CJ, Reid JG, Gunaratne PH. Expression profiling of microRNAs by deep sequencing. Briefings in bioinformatics. 2009; 10(5):490–497. [PubMed: 19332473]
- Hafner M, Renwick N, Brown M, Mihailoviæ A, Holoch D, Lin C, Pena JTG, et al. RNA-ligasedependent biases in miRNA representation in deep-sequenced small RNA cDNA libraries. RNA (New York NY). 2011; 17(9):1697–1712.
- Kawano M, Kawazu C, Lizio M, Kawaji H, Carninci P, Suzuki H, Hayashizaki Y. Reduction of noninsert sequence reads by dimer eliminator LNA oligonucleotide for small RNA deep sequencing. BioTechniques. 2010; 49(4):751–755. [PubMed: 20964636]
- Kloosterman WP, Plasterk RHA. The diverse functions of microRNAs in animal development and disease. Developmental cell. 2006; 11(4):441–450. [PubMed: 17011485]
- Lau NC, Lim LP, Weinstein EG, Bartel DP. An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science. 2001; 294(5543):858–862. [PubMed: 11679671]
- Tarasov V, Jung P, Verdoodt B, Lodygin D, Epanchintsev A, Menssen A, Meister G, et al. Differential regulation of microRNAs by p53 revealed by massively parallel sequencing: miR-34a is a p53 target that induces apoptosis and G1-arrest. Cell cycle (Georgetown, Tex.). 2007; 6(13):1586–1593.
- Thomas MF, Ansel KM. Construction of small RNA cDNA libraries for deep sequencing. Methods in molecular biology (Clifton, N.J.). 2010; 667:93–111.

- Uziel T, Karginov FV, Xie S, Parker JS, Wang Y-D, Gajjar A, He L, et al. The miR-17~92 cluster collaborates with the Sonic Hedgehog pathway in medulloblastoma. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106(8):2812–2817. [PubMed: 19196975]
- Vigneault F, Sismour AM, Church GM. Efficient microRNA capture and bar-coding via enzymatic oligonucleotide adenylation. Nature methods. 2008; 5(9):777–779. [PubMed: 19160512]
- Zhu JY, Pfuhl T, Motsch N, Barth S, Nicholls J, Grässer F, Meister G. Identification of novel Epstein-Barr virus microRNA genes from nasopharyngeal carcinomas. Journal of virology. 2009; 83(7): 3333–3341. [PubMed: 19144710]
- Zhu Q-H, Spriggs A, Matthew L, Fan L, Kennedy G, Gubler F, Helliwell C. A diverse set of microRNAs and microRNA-like small RNAs in developing rice grains. Genome research. 2008; 18(9):1456–1465. [PubMed: 18687877]

Vigneault et al. Page 10



Gel extraction, QC & Sequencing

#### **Figure 1.**

Schematic Illustration of the miRNA library preparation steps. 3'adapter is first ligated to the total RNA (steps 1–5) followed by RT primer annealing (steps 6–7) and 5'adapter ligation (steps 8–10 ). The sample is then revere transcribed (steps 11–12) and PCR enriched using bar-coded primers (steps 13–27) and is then ready for gel extraction of the miRNA library fraction and QC (steps 28–36). The library is now ready for high-throughput sequencing using either a single pass or custom indexing sequencing.



#### **Figure 2.**

Final miRNA library analyzed on 2% agarose gel. Following the PCR and clean up steps, half of the final miRNA is loaded on one lane of a 2% E-gel alongside 100 bp (lane 1) and 25 bp ladders (lane 2). The fraction corresponding to miRNA library to be sequenced was extracted on a previous gel by cutting between 125 bp and 175 bp (lane 4) as described in the procedures and also loaded alongside the non-extracted library as a visual reference and expected recovery yield.



#### **Figure 3.**

Adapter dimer consideration. Agarosegel comparing miRNA library preparation with RT primer annealing after ligation of both 3' and 5' adapter as described previously<sup>5</sup> (lane 3) against miRNA library preparation with RT primer annealing after ligation of 3' adapter only, as described here (lane 4). The adapter dimer expected to migrate at 114 bp is shown against the 100 bp (lane 1) and 25 bp ladder (lane 2). The adapter dimer is caused by ligation of 5'adapter directly to the excess of 3'adapter without miRNA captured. Our current protocol prevent the formation of this adapter dimer by annealing the RT primer to any available 3'adapter, effectively making them double stranded and therefore a poor substrate for subsequent 5'adapter ligation by T4 RNA Ligase 1.

#### **Table 1**

#### List of Oligonucleotides:



Notes: All oligonucleotides can be ordered through Integrated DNA Technologies (IDT; [http://www.idtdna.com\)](http://www.idtdna.com), with HPLC purification. The barcodes are designed to be read in a single pass read, but a custom indexing primer can also be used if desired. The adenylated adapter can be ordered from IDT, or if on a budget, made as previously described (Vigneault, Sismour, & Church, 2008).