Deep Small RNA Sequencing of *BRAF* V600E Mutated Papillary Thyroid Carcinoma With Lymph Node Metastasis

Azliana Mohamad Yusof\(^1\), Rahman Jamal\(^2\), Sazuita Saidin\(^2\), Rohaizak Muhammad\(^3\), Shahrun Niza Abdullah Suhaimi\(^3\), Isa Mohamed Rose\(^4\), Wan Fahmi Wan Nazarie\(^2\), Francis Tieng Yew Fu\(^5\) and Nurul-Syakima Ab Mutalib\(^2^*\)

\(^1\) Cytogenetics and Molecular Diagnostics Laboratory, Pantai Premier Pathology Sdn Bhd, Kuala Lumpur, Malaysia, \(^2\) UKM Medical Molecular Biology Institute (JMBi), Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia, \(^3\) Department of Surgery, Faculty of Medicine, Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia, \(^4\) Department of Pathology, Faculty of Medicine, Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia, \(^5\) Institute of Bioscience (IBS), Universiti Putra Malaysia, UPM Serdang, Selangor, Malaysia

**Keywords:** small RNA sequencing, papillary thyroid cancer, lymph node metastasis, *BRAF* V600E, microRNAs, Piwi-interacting RNA

**INTRODUCTION**

Papillary thyroid carcinoma (PTC) is the most predominant subtype of thyroid cancer, contributing to more than 80% of all thyroid or endocrine malignancies. Its prognosis is relatively good compared to other cancers, with more than 90% overall 10-year survival rate (Haugen et al., 2016). However, despite the favorable prognosis, some cases exhibit aggressive phenotype. On average, 50% of the patients are presented with lymph node metastases (LNM) at diagnosis (Saiselet et al., 2015). LNM in PTC confers various poor prognostic indicators; it increases recurrence risk and decreases long-term survival predominantly in patients older than 45 years old (Lundgren et al., 2006; Xing, 2013). Treating recurrent PTCs is still a challenge (Pasternak and Shen, 2017) and one of the important issues yet to be solved in PTC patient management is the mortality and morbidity associated with the recurrent disease (Lee et al., 2013).

MicroRNAs (miRNAs) are short, endogenous noncoding RNAs first identified in *Caenorhabditis elegans* which regulate gene expression by binding to the 3′-UTR of target messenger RNA (mRNAs). Their post-transcriptional regulatory functions are involved in controlling the level of proteins involved in numerous biological processes, including embryogenesis, organogenesis, tissue homeostasis, immune system function and cell cycle control (Tafrihi and Hasheminasab, 2019). Relationship between miRNAs and tumor growth, tumor progression and metastasis has been demonstrated by many studies, indicating the ability of these molecules to be used as biomarkers for diagnosis and prognosis (Iorio and Croce, 2012). MiRNA also plays a role as biomarker in predicting lymph nodes metastasis (LNM) in thyroid cancer (Ab Mutalib et al., 2016; Mutalib et al., 2016; Mohamad Yusof et al., 2018).

P-element-induced wimpy testis (PIWI)-interacting RNAs (piRNAs) are a quite recently discovered class of RNAs and more than 30,000 piRNA genes been identified within the human genome (Zhang et al., 2014; Wang et al., 2019). Originally described as key regulators for the germline maintenance and transposon silencing, these short RNAs were disregarded for a long time due to our limited knowledge regarding their function (Weng et al., 2019). However, emergent of new data reveals that unusual expression of piRNAs is a distinct feature in many diseases, including...
cancers (Weng et al., 2019). This has changed our perspectives on their significance in various diseases. Epigenetically, piRNAs are also the post-transcriptional regulators of the expression of specific downstream target genes (Watanabe and Lin, 2014), and accumulating evidence demonstrates that similarly to miRNAs, piRNAs also possess both oncogenic and tumor-suppressive roles (Weng et al., 2019).

BRAF is a serine-threonine kinase that is activated by RAS binding and protein recruitment to the cell membrane. Activation of MEK along with the MAPK signaling pathway is activated by BRAF phosphorylation (Nikiforov and Nikiforova, 2011). The most frequent genetic changes in PTC are point mutations of BRAF which are observed in 35 to 70% of PTC cases (Li et al., 2012; Xing, 2013). More than 95% of BRAF mutations in thyroid cancers are thymine to adenine transversion at position 1799 (T1799A) which result in the substitution of glutamate from valine at residue 600 (V600E) (Xing et al., 2015). Various studies have also shown that the BRAF V600E mutation is related to LNM, as further presented by a meta-analysis by Song et al. (2018).

In the past years, the evolution of next-generation sequencing technologies has enabled global transcriptome expression profiling and the discovery of novel human miRNAs (Kozomara and Griffiths-Jones, 2014; Kozomara and Griffiths-Jones, 2014; Kozomara et al., 2019) and piRNAs (Weng et al., 2019). Coincidentally, PTC has a relatively lower overall mutation burden hence termed as quiet genome (Cancer Genome Atlas Research Network, 2014; Siraj et al., 2016), therefore investigating the regulatory molecules such as the miRNAs and piRNAs has the potential to deepen our understanding of this cancer. This Data Report aims to provide the readers a comprehensive dataset of small RNAs expression derived from next-generation sequencing in an unbiased manner. For detailed analysis of this dataset focusing on miRNAs with biological insights in PTC, please refer to our original article published in Frontiers of Endocrinology (Mohamad Yusof et al., 2018).

**METHODS**

**Clinical Specimens**

Five pairs of tumor-adjacent normal fresh frozen thyroid tissues (n = 10) were collected from PTC patients with lymph node metastasis (N stage = N1, N1a or N1b) from the UKM Medical Centre (UKMMC). Only the thyroid specimens were collected and there was no metastatic lymph node included for several reasons. Firstly, we would like to study the source of the metastasis, which is from the cancer cells at the primary origin (the thyroid itself). Another reason was that the metastatic lymph nodes were unavailable from our Biobank. The resected lymph nodes were all sent to the Pathology Laboratory for diagnostic confirmation. Henceforth, the nodes are fixed in formalin which renders the specimens unsuitable for next-generation sequencing experiment due to RNA degradation. This study was approved by the Universiti Kebangsaan Malaysia Research Ethics Committee (UKMREC; UKM 1.5.3.5/244/UMBI-2015-002). Informed consent was obtained from all the study participants. The tissues were dissected, snap-frozen and stored in liquid nitrogen. All samples were cryosectioned and stained with hematoxylin and eosin and the percentage of tumor cells and normal cells contents were confirmed by a pathologist. The representative figures of the staining at 200× magnification were provided in **Figure 1A** (left: PTC, right: normal adjacent thyroid). Only tumor samples with at least 80% cancerous cells and normal adjacent thyroid tissues with less than 20% necrosis were selected for small RNA sequencing. The tissues were subjected to nucleic acids extraction using Allprep DNA/RNA/miRNA Universal Kit (Qiagen, Valencia, CA, USA) according to manufacturer's recommendations. The integrity of RNA was assessed using Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA) while the quantity and purity of RNA were assessed using Qubit 2.0 fluorometer and Nanodrop 2000c Spectrometer (Thermo Fisher Scientific, Pittsburgh, PA, USA), respectively.

**BRAF V600E Genotyping**

BRAF V600E genotyping was performed using forward primer 5′-TGCTTGCTCTGATAGGAAAAAG-3′ and reverse primer 5′-AGCATCTCAGGGGGCAAAAT-3′ (Monti et al., 2015). The DNA sequencing was performed on ABI Prism 3130xl Genetic Analyzer (Applied Biosystem, USA).

**Libraries Preparation**

RNA from tumor and their adjacent normal tissues were processed into libraries using TruSeq Small RNA Sample Prep Kit (Illumina, San Diego, CA, USA). Briefly, 3’ and 5’ adapters were sequentially ligated to the ends of small RNAs fractionated from 2 μg of total RNA, and reverse transcribed to generate cDNA. The cDNA was amplified using a common primer complementary to the 3’ adapter, and a primer containing 1 of 48 index sequences. Samples were size-selected (140–160 bp fragments) on a 6% polyacrylamide gel, purified, quantified and pooled for multiplexed sequencing.

**Small RNA Sequencing**

The resulting pooled libraries were normalized to 2 nM and were hybridized to oligonucleotide-coated single-read flow cells for cluster generation using HiSeq® Rapid SR Cluster Kit v2 on Hiseq 2500 (Illumina, San Diego, CA, USA). Subsequently, the clustered pooled microRNA libraries were sequenced on the HiSeq 2500 for 50 sequencing cycles using HiSeq® Rapid SBS Kit v2 (50 Cycle) (Illumina, San Diego, CA, USA). The sequencing run was completed on December 18, 2015.

**Bioinformatics and Statistical Analyses**

Pre- and post-processing of data were executed in BaseSpace software (Illumina, San Diego, CA, USA). Illumina Sequence Integration Software (Isis) version 2.5.52.11 processed raw data from Illumina sequencers into FASTQ files. Subsequently, Small RNA Analysis app version 1.0.0.0 was used for determination of differentially expressed small RNAs. Differential expression analysis is performed for the following marker types; miRNA family, precursor group, miRNAs (include mature miRNAs and isomiRs) and piRNAs. Briefly, this workflow uses Bowtie version
0.12.8 to align each cluster against abundant, mature miRNA, other RNA, and genomic reference databases (Langmead et al., 2009). The reference genome used for the alignment was Homo sapiens maskedPAR (UCSC hg19). The output from Bowtie was then processed by SAMtools 0.1.19-isis-1.0.2 (Li et al., 2009), annotated to miRBase v21 (Kozomara and Griffiths-Jones, 2014; Kozomara et al., 2019) and performed differential analysis using DESeq2 (Love et al., 2014). DESeq2 removed low expressed markers with normalized mean count < 10 before testing. The DESeq2 variance model is used to detect and exclude outliers based on the extreme variation between replicates. The status (filtered or passed) and the result of the analysis (mean expression, fold change, standard error, p-value, q-value, etc.) were reported for each marker and will be available upon request. Additionally,
There were 33 upregulated and 25 downregulated precursor miRNAs, respectively.

**Novel Precursor miRNAs**

There were 95 novel precursors predicted from this dataset, with 36 novel precursors passed the filter and subjected to differential analysis. A set of 16 novel precursor miRNAs were significantly differentially expressed (q value < 0.05) in cancer versus normal thyroid tissues. There were 4 upregulated and 12 downregulated novel precursor miRNAs, respectively.

**miRNA Family**

Expression profiles for 418 previously annotated miRNA families were detected and 150 families passed the filter and subjected to differential analysis. A set of 36 miRNA families were significantly differentially expressed (q value < 0.05) in cancer versus normal thyroid tissues. There were 17 upregulated and 19 downregulated miRNA families, respectively. The workflow does not predict novel miRNA families.

**piRNA**

Expression profiles for 848 previously annotated piRNAs were detected and 110 piRNAs passed the filter and subjected to differential analysis. A set of 19 piRNAs were significantly differentially expressed (q value < 0.05) in cancer versus normal thyroid tissues. There were 8 upregulated piRNAs and 11 downregulated piRNAs. The workflow does not predict novel piRNAs.

**Direct Link to Deposited Data and Information to Users**

The small RNA sequencing data in fastq format was deposited at the NCBI Sequence Read Archive (SRA) at http://www.ncbi.nlm.nih.gov/sra with accession number PRJNA378313 (SRP101463). The sample IDs are TY135, TY 146, TY2, TY5 and TY5 and

---

**TABLE 1** | Summary of small RNA sequencing results.

|                   | Known | Novel | Total |
|-------------------|-------|-------|-------|
| miRNAs            |       |       |       |
| Total miRNAs      | 4,662 | 173   | 4,835 |
| Tested miRNAs     | 1,079 | 26    | 1,105 |
| Differentially expressed miRNAs | 252 | 16 | 268 |
| Precursors        |       |       |       |
| Total Precursor Groups | 971 | 95 | 1,066 |
| Tested Precursor Groups | 219 | 36 | 255 |
| Differentially expressed Precursor Groups | 58 | 16 | 74 |
| miRNA families    |       |       |       |
| Total miRNA Families | 418 | 0 | 418 |
| Tested miRNA Families | 150 | 0 | 150 |
| Differentially expressed miRNA families | 36 | 0 | 36 |
| piRNAs            |       |       |       |
| Total piRNAs      | 848   | 0     | 848   |
| Tested piRNAs     | 110   | 0     | 110   |
| Differentially expressed piRNAs | 19 | 0 | 19 |

Tested markers are those that pass the low-count filter and are tested for statistical significance. Novel markers (when applicable) are predicted by miRDeep*.
TY60. This raw data will enable other researchers to perform mapping and differential expression analysis according to their specific objectives.

**DATA AVAILABILITY STATEMENT**

The datasets generated for this study can be found in the NCBI Sequence Read Archive (SRA), PRJNA378313 (SRP101463).

**ETHICS STATEMENT**

The studies involving human participants were reviewed and approved by Universiti Kebangsaan Malaysia Research Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

**AUTHOR CONTRIBUTIONS**

AMY and N-SAM involved in the specimen collections, libraries preparation and sequencing, data analyses, acquisition of data and drafting of the manuscript. RJ provides a critical review of the manuscript. SS performed BRAF V600E genotyping, cryosectioning and staining of the fresh frozen tissues. RM and SNAS are thyroid surgeons involved in specimen retrieval. IMR assessed tumor percentage of the tissues. FTYF was involved in the initial draft. WFHN is the bioinformatician partly involved in data analysis. All authors read and approved the final manuscript.

**REFERENCES**

Ab Mutalib, N.-S., Othman, S. N., Mohamad Yusof, A., Abdullah Suhaimi, S. N., Muhammad, R., and Jamal, R. (2016). Integrated microRNA gene expression and transcription factors signature in papillary thyroid cancer with lymph node metastasis. *PeerJ* 4, e2119. doi:10.7717/peerj.2119.

An, J., Lai, J., Lehman, M. L., and Nelson, C. C. (2013). miRDeep*: an integrated application tool for miRNA identification from RNA sequencing data. *Nucleic Acids Res.* 41, 727–739. doi:10.1093/nar/gkt1181.

Lee, J. C., Zhao, J. T., Clifton-Bligh, R. J., Gill, A., Gundara, J. S., Ip, J. C., et al. (2013). microRNA involvement in human cancer. *Genome Biol.* 14, 569–580. doi:10.1186/guk120.

Iorio, M. V., and Croce, C. M. (2012). microRNA involvement in human cancer. *Carcinogenesis* 33, 1126–1133. doi:10.1093/carcin/bgs140.

Kozomara, A., Birgaonanu, M., and Griffiths-Jones, S. (2019). miRBase: from microRNA sequences to function. *Nucleic Acids Res.* 47, D155–D162. doi:10.1093/nar/gky1141.

Kozomara, A., and Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* 42, D68–D73. doi:10.1093/nar/gkt1181.

Kozomara, A., and Griffiths-Jones, S. (2011). miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* 39, D152–D157. doi:10.1093/nar/gkq1027.

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, R25. doi:10.1186/gb-2009-10-3-r25.

Lee, J. C., Zhao, J. T., Clifton-Bligh, R. J., Gill, A., Gundara, J. S., Ip, J. C., et al. (2013). MicroRNA-222 and microRNA-146b are tissue and circulating biomarkers of recurrent papillary thyroid cancer. *Cancer* 119, 4358–4365. doi:10.1002/cncr.28254.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinfroma.* *Oxf. Engl.* 25, 2078–2079. doi:10.1093/bioinformatics/btp352.

**FUNDING**

This research was funded by the Fundamental Research Grant Scheme (FRGS) from the Ministry of Education Malaysia (FRGS/1/2014/SKK01/UKM/03/1).

**ACKNOWLEDGMENTS**

The authors thanks Miss Chin Minning, the former Field Application Specialist from Science Vision Sdn Bhd for her extensive support during library preparation and next-generation sequencing.

Li, X., Abdel-Mageed, A. B., and Kandil, E. (2012). BRAF mutation in papillary thyroid carcinoma. *Int. J. Clin. Exp. Med.* 5, 310–315.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi:10.1186/s13059-014-0550-8.

Monti, E., Bovero, M., Mortara, L., Pera, G., Zupo, S., Gugliatti, E., et al. (2015). BRAF mutations in an Italian regional population: implications for the therapy of thyroid cancer. *Int. J. Endocrinol.* 2015, 138734. doi:10.1155/2015/138734.

Kozomara, A., and Griffiths-Jones, S. (2011). miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* 39, D152–D157. doi:10.1093/nar/gkq1027.

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, R25. doi:10.1186/gb-2009-10-3-r25.

Lee, J. C., Zhao, J. T., Clifton-Bligh, R. J., Gill, A., Gundara, J. S., Ip, J. C., et al. (2013). MicroRNA-222 and microRNA-146b are tissue and circulating biomarkers of recurrent papillary thyroid cancer. *Cancer* 119, 4358–4365. doi:10.1002/cncr.28254.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinfroma.* *Oxf. Engl.* 25, 2078–2079. doi:10.1093/bioinformatics/btp352.

**ACKNOWLEDGMENTS**

The authors thanks Miss Chin Minning, the former Field Application Specialist from Science Vision Sdn Bhd for her extensive support during library preparation and next-generation sequencing.

Li, X., Abdel-Mageed, A. B., and Kandil, E. (2012). BRAF mutation in papillary thyroid carcinoma. *Int. J. Clin. Exp. Med.* 5, 310–315.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi:10.1186/s13059-014-0550-8.

Monti, E., Bovero, M., Mortara, L., Pera, G., Zupo, S., Gugliatti, E., et al. (2015). BRAF mutations in an Italian regional population: implications for the therapy of thyroid cancer. *Int. J. Endocrinol.* 2015, 138734. doi:10.1155/2015/138734.

Kozomara, A., and Griffiths-Jones, S. (2011). miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* 39, D152–D157. doi:10.1093/nar/gkq1027.

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, R25. doi:10.1186/gb-2009-10-3-r25.

Lee, J. C., Zhao, J. T., Clifton-Bligh, R. J., Gill, A., Gundara, J. S., Ip, J. C., et al. (2013). MicroRNA-222 and microRNA-146b are tissue and circulating biomarkers of recurrent papillary thyroid cancer. *Cancer* 119, 4358–4365. doi:10.1002/cncr.28254.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinfroma.* *Oxf. Engl.* 25, 2078–2079. doi:10.1093/bioinformatics/btp352.
Song, J.-Y., Sun, S.-R., Dong, F., Huang, T., Wu, B., and Zhou, J. (2018). Predictive value of BRAFV600E mutation for lymph node metastasis in papillary thyroid cancer: a meta-analysis. *Curr. Med. Sci.* 38, 785–797. doi: 10.1007/s11596-018-1945-7.

Tafrihi, M., and Hasheminasab, E. (2019). MiRNAs: biology, biogenesis, their web-based tools, and databases. *MicroRNA Sharigah United Arab Emir.* 8, 4–27. doi: 10.2174/2211536607666180827111633.

Wang, J., Zhang, P., Lu, Y., Li, Y., Zheng, Y., Kan, Y., et al. (2019). piRBase: a comprehensive database of piRNA sequences. *Nucleic Acids Res.* 47, D175–D180. doi: 10.1093/nar/gky1043.

Watanabe, T., and Lin, H. (2014). Posttranscriptional regulation of gene expression by Piwi proteins and piRNAs. *Mol. Cell* 56, 18–27. doi: 10.1016/j.molcel.2014.09.012.

Weng, W., Li, H., and Goel, A. (2019). Piwi-interacting RNAs (piRNAs) and cancer: emerging biological concepts and potential clinical implications. *Biochim. Biophys. Acta Rev. Cancer* 1871, 160–169. doi: 10.1016/j.bbcan.2018.12.005.

Xing, M. (2013). Molecular pathogenesis and mechanisms of thyroid cancer. *Nat. Rev. Cancer* 13, 184–199. doi: 10.1038/nrc3431.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Yusof, Jamal, Muhammad, Suhaimi, Rose, Nazarie, Fu and Mutalib. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.