



Previously undescribed thyroid-specific miRNA sequences in papillary thyroid carcinoma

Mateus Camargo Barros-Filho^{1,2} · Michelle Pewarchuk¹ · Brenda de Carvalho Minatel¹ · Adam P. Sage¹ · Erin A. Marshall¹ · Victor D. Martinez¹ · Leigha D. Rock^{1,3,4} · Gavin MacAulay¹ · Luiz Paulo Kowalski² · Silvia Regina Rogatto⁵ · Cathie Garnis¹ · Wan L. Lam¹

Received: 6 December 2018 / Revised: 4 February 2019 / Accepted: 11 February 2019 / Published online: 6 March 2019
© The Author(s), under exclusive licence to The Japan Society of Human Genetics 2019

Abstract

Papillary thyroid carcinoma (PTC) is the most common thyroid malignancy, wherein diagnostic limitations and lack of accurate prognostic factors are important clinical challenges. In this study, we report the discovery of 234 novel miRNAs in non-neoplastic thyroid and PTC samples, obtained from publicly available small RNA sequencing datasets (TCGA and GEO). These sequences were observed to display similar molecular features compared to currently annotated miRNAs. These potentially novel miRNAs presented tissue-specificity and largely decreased expression in PTC compared to non-neoplastic samples. We showed that the disrupted novel miRNAs have diagnostic and prognostic potential, and were associated with *BRAF* mutation, a frequent alteration related to more aggressive PTC. In conclusion, our results expand the miRNA repertoire in thyroid tissues and highlight the potential biological role and clinical utility of previously unannotated miRNAs.

Advances in imaging techniques have led to a drastic increase in the detection of thyroid nodules; however, ~90% of these nodules are benign [1]. Due to diagnostic limitations, patients with benign thyroid disease are frequently subjected to unnecessary surgeries [1]. In addition, only a

small fraction of papillary thyroid carcinoma (PTC), the most prevalent thyroid malignancy, is associated with poor outcome [2]. Hence, there is an urgent need to expand the repertoire of specific PTC markers, which can potentially aid in the management of thyroid cancer [1, 3, 4].

MicroRNAs (miRNAs) are frequently disrupted in many cancer types and are suitable candidates to be used as biomarkers [5]. Recent large-scale sequencing has allowed for the discovery of novel miRNAs, which may have greater clinical utility than their well-characterized and widely expressed counterparts [6–8]. The presence of these overlooked miRNAs is explained by the lack of stable expression across different tissues and low-coverage techniques used in early studies [6–8]. Thus, we sought to identify previously unannotated miRNAs expressed in and specific to thyroid tissues.

We applied our custom discovery pipeline to identify previously unannotated miRNAs from 504 PTC and 59 adjacent non-malignant thyroid tissues (NT) processed by The Cancer Genome Atlas (TCGA) (Figure S1 and Table S1) (<https://portal.gdc.cancer.gov/>) [7]. Using an algorithm analogous to miRDeep2 [9] publicly available through the miRMaster online tool [10], we generated a broad list of novel miRNA candidates [10] (Supplementary Methods and Table S2). We curated this list based on the

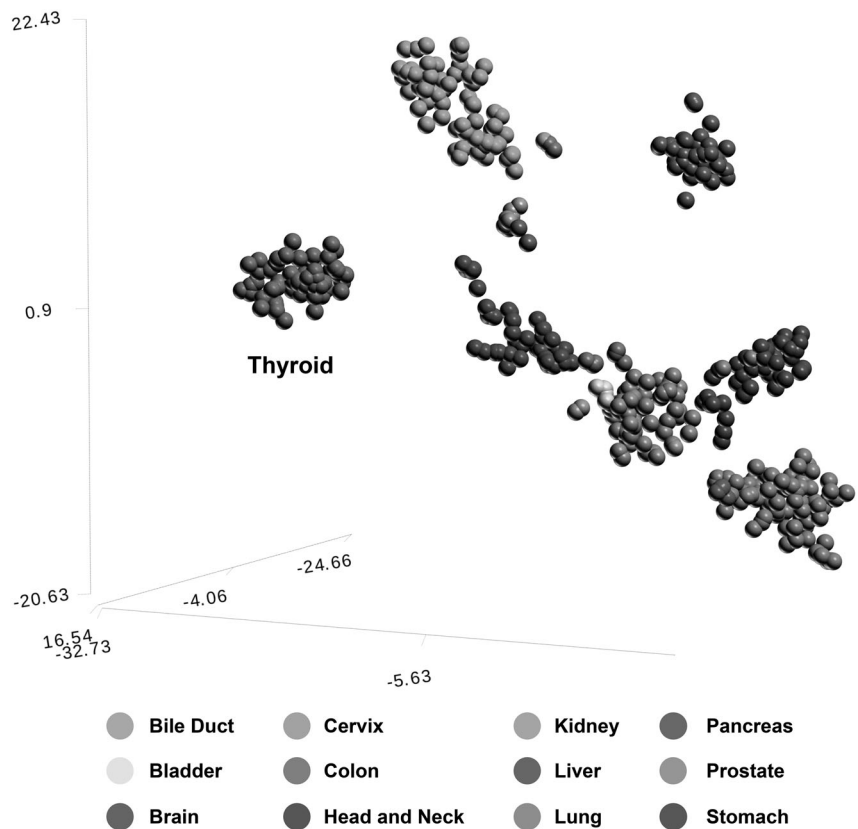
These authors contributed equally: Mateus Camargo Barros-Filho, Michelle Pewarchuk, Brenda de Carvalho Minatel

Supplementary information The online version of this article (<https://doi.org/10.1038/s10038-019-0583-7>) contains supplementary material, which is available to authorized users.

✉ Mateus Camargo Barros-Filho
mbarros@bccrc.ca

- ¹ Department of Integrative Oncology, British Columbia Cancer Research Centre, Vancouver, Canada
- ² International Research Center, A.C. Camargo Cancer Center, Sao Paulo, Brazil
- ³ Department of Cancer Control Research, British Columbia Cancer Research Centre, Vancouver, Canada
- ⁴ Faculty of Dentistry, University of British Columbia, Vancouver, Canada
- ⁵ Vejle Hospital, Institute of Regional Health Research, University of Southern Denmark, Odense, Denmark

Fig. 1 t-Distributed Stochastic Neighbor Embedding (t-SNE) analysis of the novel miRNAs in non-neoplastic tissues from TCGA. A suggestive thyroid-specific expression pattern of the 217 Tnov-miRs was detected in non-malignant thyroid tissues (RPM > 1 in more than 10% of the NT group). The distribution of the samples in the three-dimensional chart demonstrates a cluster of thyroid samples (red), clearly separated from 12 other non-malignant tissue types from TCGA (bile duct, bladder, brain, cervix, colon, head and neck, kidney, liver, lung, pancreas, prostate and stomach)



expression (RPM ≥ 1 in 10% of samples) and molecular features of these transcripts, resulting in a robust set of 234 previously unannotated miRNAs in thyroid samples (herein referred to as Tnov-miRs). A total of 92 sequences were exclusively expressed in non-malignant samples, 17 in tumors, and 125 shared between both sample types.

The expression range of the Tnov-miRs is significantly ($p < 0.001$) different from currently annotated sequences (Figure S2). Despite the differences in expression levels, the Tnov-miRs share genomic features characteristic of annotated miRNAs such as their widespread genomic localization and distribution (Figure S3), seed sequence nucleotide composition (positions 2–7), GC content, as well as secondary structure and canonical processing by DICER (Figures S4, 5 Table S3). To further exemplify the miRNA-like characteristics of these novel sequences, we found eight Tnov-miRs that share seed sequences with miRNAs derived from the miR-17 family, which is frequently deregulated in numerous cancers [11, 12]. As expected, these Tnov-miRs were also predicted to regulate a similar set of target genes (Figure S5). Collectively, these results support the characterization of these novel transcripts as miRNAs.

Tnov-miRs were observed to be thyroid-specific, as evidenced by t-Distributed Stochastic Neighbor Embedding (t-SNE) analysis. Their combined expression pattern in each tissue was able to clearly distinguish non-malignant thyroid

samples from other tissue types (Fig. 1). This tissue-specific nature highlights their potential relevance to thyroid biology.

Interestingly, 125 of the 234 novel miRNAs detected were differentially expressed (BH- $p < 0.001$, fold-change > 1.5) between matched tumor and non-malignant thyroid samples ($n = 59$ pairs). Most sequences ($n = 115$) were found to be significantly under-expressed in PTC, while only a fraction ($n = 10$) showed over-expression (Table S4). Although miRNAs are known to have either tumor-suppressive or oncogenic functions, an overall downregulation of their expression has been described as a common feature of tumorigenesis through the impairment of miRNA biogenesis [13, 14]. A recent study demonstrated that restoration of *DICER1* expression inhibits PTC invasion and metastasis, suggesting that the loss of functional miRNAs may be important in the development and progression of PTC [15].

To validate the expression of these Tnov-miRs, we analyzed an independent dataset of 17 PTC and eight NT samples using the same criteria (GSE63511). Despite the limited sample size of this dataset, 16 of our Tnov-miRs were predicted and nine met the expression threshold criteria, confirming the existence of these sequences. Among them, Tnov-mir-74 and Tnov-mir-161 were over-expressed and Tnov-mir-231 under-expressed in PTC (Fig. 2a and Figure S6). Further, the combined expression pattern of these

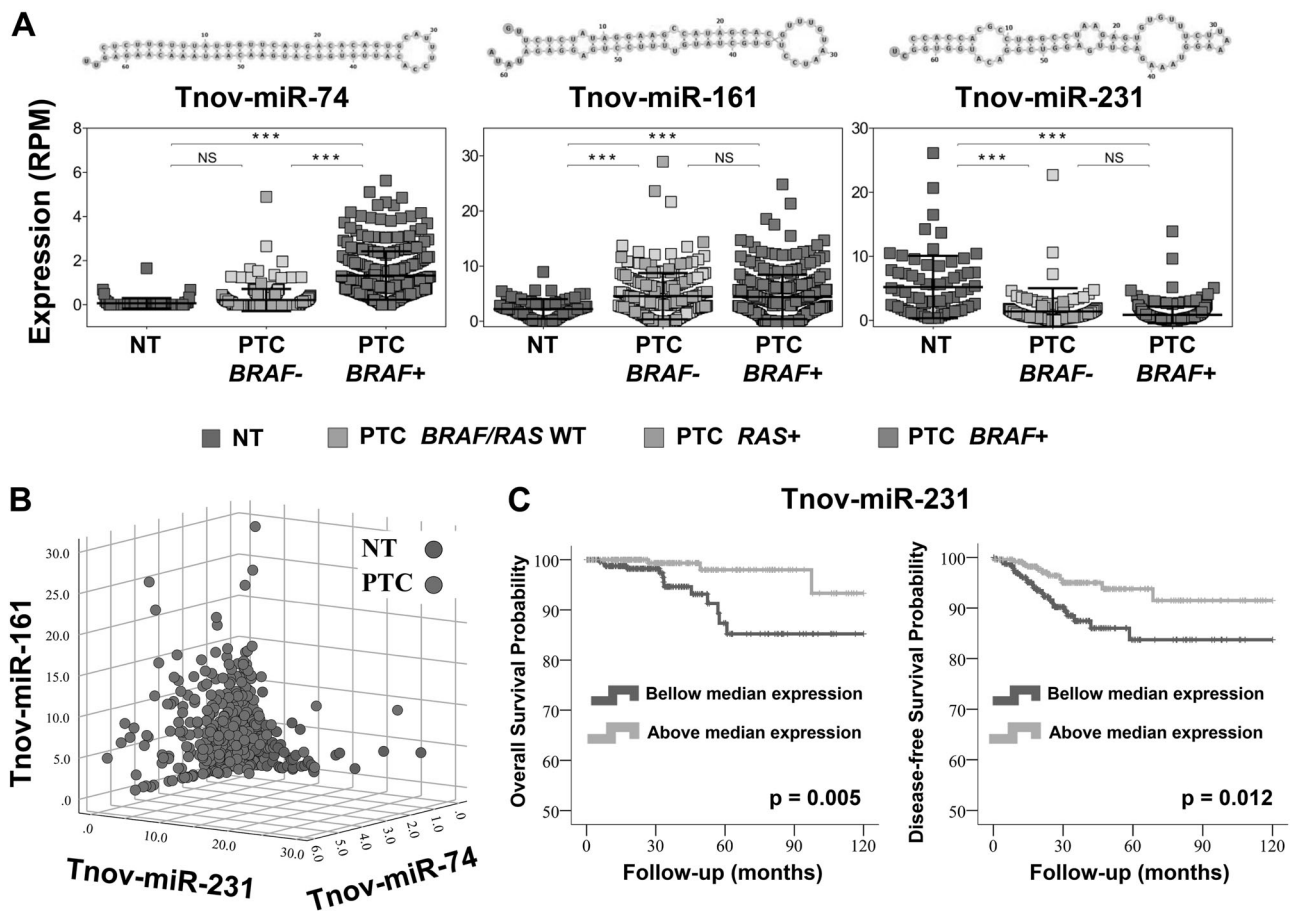


Fig. 2 Potential clinical applications of the disrupted novel miRNAs. **a** Visual representation of the structure of the novel miRNA precursors generated by the novoMiRank tool and expression of Tnov-mir-74, Tnov-mir-161, and Tnov-mir-231 among non-malignant (blue), papillary thyroid carcinoma negative for *BRAF* mutation (gray and green) and with *BRAF* mutation (red) samples from the TCGA cohort. **b** Three-dimensional scatter-plot comprising the expression of the three novel miRNAs, demonstrating their potential diagnostic value

Tnov-miRs was capable of correctly classifying (Diagonal Linear Discriminant Analysis classifier) tumors from NT (TCGA cohort) with 91% accuracy (94.1% sensitivity, 66.1% specificity; Fig. 2b). The three-novel miRNA model presented superior performance than the use of single markers (Figure S7). These findings pointed a potential role of these Tnov-miRs in the diagnosis of thyroid lesions.

We verified that Tnov-mir-74 increased expression was exclusively observed in *BRAF*V600E-PTC, a frequently observed genetic event and related to more aggressive PTC cases [2, 16, 17] (Fig. 2a). Tnov-mir-74 coding region was found hypomethylated in PTC compared to NT samples ($\Delta\beta = -0.31$; BH- $p < 0.001$) and negatively correlated with miRNA expression ($r = -0.605$; $p < 0.001$). The same genomic region (probe-ID: cg25571414) had been previously reported to be hypomethylated in *BRAF*V600E compared to *BRAF*WT-PTC [18]. In fact, global

hypomethylation has been shown to be induced by *BRAF* mutation [18], which can increase the expression of specific miRNAs [19]. In addition, the decreased expression of Tnov-mir-231 was found to be associated with lower overall and disease-free survival (Fig. 2c). We observed that the region encoding Tnov-mir-234 (22q13.2) presented hemizygous deletion in 18% of the PTC samples; however, no association was observed with its decreased expression ($p = 0.715$).

due to the separation of PTC (red) from NT samples (blue). **c** Kaplan–Meier representation (ranging from 50 to 100% survival probability) of overall and disease-free survival from PTC patients according to Tnov-mir-231, demonstrating a worse prognosis in patients presenting lower levels of the predicted miRNA (P from log rank test). NS not significant; *** $P < 0.001$ (ANOVA followed by Tukey post-hoc); PTC papillary thyroid carcinoma, NT non-malignant thyroid tissue

In order to predict their possible biological roles, a miRNA–mRNA integrative analysis was carried out with Tnov-miRs and their predicted mRNA targets (miRanda algorithm). Negative correlations ($p < 0.05$) were observed between Tnov-mir-74, Tnov-mir-161, and Tnov-mir-234 and 36, 69, and 212 mRNAs, respectively (Table S5). A subsequent in-silico analysis (pathDIP, BH- $p < 0.05$) revealed an enrichment for cancer-related pathways (Table S6). These

results highlight a potential biological role of the Tnov-miRs in thyroid cancer onset.

In conclusion, we identified 234 Tnov-miRs expressed in thyroid tissues with potential relevance to PTC biology. While our study was performed on a predictive platform and further validation will be required, we provide an additional resource for further exploration of these transcripts. Our results suggest that the incorporation of previously unannotated miRNAs in the development of diagnostic and prognostic stratification panels may increase their accuracy.

Acknowledgements This work was supported by grants from the Canadian Institutes for Health Research (CIHR FDN-143345) and the São Paulo Research Foundation (FAPESP 2015/20748-5). MCBF are supported by scholarships from FAPESP (2015/17707-5 and 2018/06138-8). BCM, APS, EAM and VDM are supported by scholarships from the University of British Columbia. APS, EAM and LDR are also supported by scholarships from CIHR. EAM is a Vanier Canada Graduate Scholar.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Durante C, Grani G, Lamartina L, Filetti S, Mandel SJ, Cooper DS. The diagnosis and management of thyroid nodules: a review. *JAMA*. 2018;319:914–24.
- Lamartina L, Grani G, Durante C, Borget I, Filetti S, Schlumberger M. Follow-up of differentiated thyroid cancer - what should (and what should not) be done. *Nat Rev Endocrinol*. 2018;14:538–51.
- Fagin JA, Wells SA. Biologic and clinical perspectives on thyroid cancer. *N Engl J Med*. 2016;375:1054–67.
- Eszlinger M, Lau L, Ghaznavi S, Symonds C, Chandarana SP, Khalil M, et al. Molecular profiling of thyroid nodule fine-needle aspiration cytology. *Nat Rev Endocrinol*. 2017;13:415–24.
- Gebert LFR, MacRae IJ. Regulation of microRNA function in animals. *Nat Rev Mol Cell Biol*. 2019;20:21–37.
- Minatel BC, Martinez VD, Ng KW, Sage AP, Tokar T, Marshall EA, et al. Large-scale discovery of previously undetected microRNAs specific to human liver. *Hum Genomics*. 2018;12:16.
- Sage AP, Minatel BC, Marshall EA, Martinez VD, Stewart GL, Enfield KSS, et al. Expanding the miRNA transcriptome of human kidney and renal cell carcinoma. *Int J Genom*. 2018;2018:6972397.
- Londin E, Loher P, Telonis AG, Quann K, Clark P, Jing Y, et al. Analysis of 13 cell types reveals evidence for the expression of numerous novel primate- and tissue-specific microRNAs. *Proc Natl Acad Sci USA*. 2015;112:E1106–15.
- An J, Lai J, Lehman ML, Nelson CC. miRDeep*: an integrated application tool for miRNA identification from RNA sequencing data. *Nucleic Acids Res*. 2013;41:727–37.
- Fehlmann T, Backes C, Kahraman M, Haas J, Ludwig N, Posch AE, et al. Web-based NGS data analysis using miRMaster: a large-scale meta-analysis of human miRNAs. *Nucleic Acids Res*. 2017;45:8731–44.
- Mogilyansky E, Rigoutsos I. The miR-17/92 cluster: a comprehensive update on its genomics, genetics, functions and increasingly important and numerous roles in health and disease. *Cell Death Differ*. 2013;20:1603–14.
- Hamilton MP, Rajapakse K, Hartig SM, Reva B, McLellan MD, Kandath C, et al. Identification of a pan-cancer oncogenic microRNA superfamily anchored by a central core seed motif. *Nat Commun*. 2013;4:2730.
- Lin S, Gregory RI. MicroRNA biogenesis pathways in cancer. *Nat Rev Cancer*. 2015;15:321–33.
- Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, et al. MicroRNA expression profiles classify human cancers. *Nature*. 2005;435:834–8.
- Yuan X, Mu N, Wang N, Strååt K, Sofiadis A, Guo Y, et al. GABPA inhibits invasion/metastasis in papillary thyroid carcinoma by regulating DICER1 expression. *Oncogene*. 2018;38:965–79.
- Xing M, Alzahrani AS, Carson KA, Shong YK, Kim TY, Viola D, et al. Association between BRAF V600E mutation and recurrence of papillary thyroid cancer. *J Clin Oncol*. 2015;33:42–50.
- Shen X, Zhu G, Liu R, Viola D, Elisei R, Puxeddu E, et al. Patient age-associated mortality risk is differentiated by BRAF V600E status in papillary thyroid cancer. *J Clin Oncol*. 2018;36:438–45.
- Beltrami CM, Dos Reis MB, Barros-Filho MC, Marchi FA, Kuasne H, Pinto CAL, et al. Integrated data analysis reveals potential drivers and pathways disrupted by DNA methylation in papillary thyroid carcinomas. *Clin Epigenetics*. 2017;9:45.
- Ortiz IMDP, Barros-Filho MC, Dos Reis MB, Beltrami CM, Marchi FA, Kuasne H, et al. Loss of DNA methylation is related to increased expression of miR-21 and miR-146b in papillary thyroid carcinoma. *Clin Epigenetics*. 2018;10:144.