

# MicroRNA deep-sequencing reveals master regulators of follicular and papillary thyroid tumors

Veronika Mancikova<sup>1,13</sup>, Esmeralda Castelblanco<sup>2,13</sup>, Elena Pineiro-Yanez<sup>3</sup>, Javier Perales-Paton<sup>3</sup>, Aguirre A de Cubas<sup>1</sup>, Lucia Inglada-Perez<sup>1,4</sup>, Xavier Matias-Guiu<sup>5</sup>, Ismael Capel<sup>6</sup>, Maria Bella<sup>7</sup>, Enrique Lerma<sup>8</sup>, Garcilaso Riesco-Eizaguirre<sup>9,10</sup>, Pilar Santisteban<sup>9</sup>, Francisco Maravall<sup>2</sup>, Didac Mauricio<sup>11,12</sup>, Fatima Al-Shahrour<sup>3</sup> and Mercedes Robledo<sup>1,4</sup>

<sup>1</sup>Hereditary Endocrine Cancer Group, Spanish National Cancer Research Centre (CNIO), Madrid, Spain;

<sup>2</sup>Department of Endocrinology and Nutrition, University Hospital Arnau de Vilanova, IRBLLEIDA, Lleida, Spain;

<sup>3</sup>Translational Bioinformatics Unit, Clinical Research Program, Spanish National Cancer Research Centre (CNIO), Madrid, Spain; <sup>4</sup>ISCIII Center for Biomedical Research on Rare Diseases (CIBERER), Madrid, Spain;

<sup>5</sup>Department of Pathology, University Hospital Arnau de Vilanova, University of Lleida, IRBLLEIDA, Lleida, Spain; <sup>6</sup>Department of Endocrinology and Nutrition, Hospital de Sabadell, Sabadell, Barcelona, Spain;

<sup>7</sup>Department of Pathology, Hospital de Sabadell, Sabadell, Barcelona, Spain; <sup>8</sup>Department of Pathology, Hospital Santa Creu i Sant Pau, Barcelona, Spain; <sup>9</sup>Instituto de Investigaciones Biomedicas ‘Alberto Sols’, Madrid, Spain;

<sup>10</sup>Hospital Universitario de Mostoles, Móstoles, Madrid, Spain; <sup>11</sup>Germans Trias i Pujol Health Sciences Research Institute (IGTP), Badalona, Spain and <sup>12</sup>Department of Endocrinology and Nutrition, University Hospital Germans Trias i Pujol, Department of Medicine, Autonomous University of Barcelona, Barcelona, Spain

**MicroRNA deregulation could be a crucial event in thyroid carcinogenesis. However, current knowledge is based on studies that have used inherently biased methods. Thus, we aimed to define in an unbiased way a list of deregulated microRNAs in well-differentiated thyroid cancer in order to identify diagnostic and prognostic markers. We performed a microRNA deep-sequencing study using the largest well-differentiated thyroid tumor collection reported to date, comprising 127 molecularly characterized tumors with follicular or papillary patterns of growth and available clinical follow-up data, and 17 normal tissue samples. Furthermore, we integrated microRNA and gene expression data for the same tumors to propose targets for the novel molecules identified. Two main microRNA expression profiles were identified: one common for follicular-pattern tumors, and a second for papillary tumors. Follicular tumors showed a notable overexpression of several members of miR-515 family, and downregulation of the novel microRNA miR-1247. Among papillary tumors, top upregulated microRNAs were miR-146b and the miR-221~222 cluster, while miR-1179 was downregulated. *BRAF*-positive samples displayed extreme downregulation of miR-7 and -204. The identification of the predicted targets for the novel molecules gave insights into the proliferative potential of the transformed follicular cell. Finally, by integrating clinical follow-up information with microRNA expression, we propose a prediction model for disease relapse based on expression of two miRNAs (miR-192 and let-7a) and several other clinicopathological features. This comprehensive study complements the existing knowledge about deregulated microRNAs in the development of well-differentiated thyroid cancer and identifies novel markers associated with recurrence-free survival.**

*Modern Pathology* (2015) 28, 748–757; doi:10.1038/modpathol.2015.44; published online 27 February 2015

Correspondence: Dr M Robledo, PhD or Dr F Al-Shahrour, PhD, Hereditary Endocrine Cancer Group, Spanish National Cancer Research Center (CNIO), Madrid, Spain or Translational Bioinformatics Unit, Clinical Research Program, Spanish National Cancer Research Centre (CNIO), Madrid, Spain.

E-mail: mrobledo@cnio.es or falshahrour@cnio.es

<sup>13</sup>These authors contributed equally to this work.

Received 13 October 2014; revised 16 January 2015; accepted 16 January 2015; published online 27 February 2015

MicroRNAs (miRNAs) are potent regulators of gene expression in a tissue-specific manner. Their deregulation has been demonstrated to be a hallmark of cancer. It has been suggested that miRNA profiles can reliably identify the cell origin of tumors and that they are specific to differentiation stage and driver alterations, and associated with progression and response to treatment.<sup>1–3</sup> Thus, their clinical

utility has been extensively studied for many cancer types.

It is especially appealing to explore the diagnostic and prognostic value of miRNA profiles in thyroid cancer, since this is a very complex and heterogeneous disease showing various stages of differentiation. Further, the accurate diagnosis of some patients remains an unresolved challenge in the clinical setting.<sup>4</sup> The vast majority of thyroid tumors develop from follicular epithelial cells of the gland, often after the emergence of a driver mutation altering the MAPK pathway (affecting key genes such as *BRAF*, *RAS*, or *RET*). They include benign (follicular adenoma) and malignant forms (papillary thyroid carcinoma; follicular carcinoma). Both papillary and follicular thyroid carcinomas may progress to poorly differentiated thyroid carcinoma or lose their differentiation completely (anaplastic thyroid cancer), or the dedifferentiated disease can emerge spontaneously.<sup>5</sup> There is an urgent need to understand the biology underlying both progressive and dedifferentiated disease, as there is still no effective treatment available for patients diagnosed with these subtypes.<sup>6</sup> Although the remaining thyroid cancers, termed well-differentiated thyroid carcinomas, are generally indolent malignancies with well-established clinical management and excellent prognosis,<sup>7</sup> a clinically relevant proportion presents recurrent disease and it is equally important to pinpoint predictive markers of disease relapse.

MiRNA profiling has been already extensively applied to well-differentiated thyroid cancer. Similarly to other genomic features,<sup>8–10</sup> the miRNA expression fingerprints have been shown to be subtype- and driver alteration-specific (reviewed in Pallante *et al*<sup>11</sup>). Moreover, the role of polymorphisms in the complementary sites of target mRNAs was initially proposed in the context of papillary carcinomas,<sup>12</sup> and polymorphisms in the miRNA sequence itself were shown to be predisposing to thyroid cancer,<sup>13</sup> highlighting the role of miRNAs in thyroid tumorigenesis. Upregulation of miR-146b and the miR-221 ~ 222 cluster are the most commonly documented miRNA changes related to papillary tumors.<sup>12,14–17</sup> However, little data are available on downregulated miRNAs, or on the changes present in less frequent histological subtypes such as follicular tumors. It is also important to highlight that apart from recent miRNA deep-sequencing studies using papillary tumors,<sup>16,18,19</sup> the remaining data on deregulated miRNAs in thyroid cancer were generated at a time when only ~300 miRNAs had been identified, compared with the ~1300 miRNAs that are known today.<sup>11</sup>

We have performed a miRNA deep-sequencing study using the largest collection of thyroid samples published to date, comprising of 127 thyroid tumors (including 26 follicular adenomas, 23 follicular carcinomas and 78 papillary thyroid carcinomas) and 17 normal thyroid tissues. We not only confirm some of the miRNA expression changes previously described as implicated in thyroid cancer, but also identify novel ones, such as the downregulation of

miR-1247 in tumors with a follicular pattern of growth. By integrating miRNA and mRNA expression data from the same tumors, we identify several possible targets for the top novel deregulated candidates. Finally, we propose a relapse prediction model based on expression of two miRNAs and several other clinicopathological features.

## Materials and methods

### Thyroid Sample Collection and Patient Follow-Up

One hundred and twenty-seven thyroid tumors were snap frozen following surgery at *Hospital Sant Pau* and *Hospital Sabadell* in Barcelona (Spain) and at *Hospital Arnau de Vilanova* in Lleida (Spain), and stored at  $-80^{\circ}\text{C}$ . An informed consent was obtained from all the study participants, and the study was approved by the Institutional Review Board (*comité de bioética y bienestar animal*) of the *Instituto de Salud Carlos III*. Of the samples collected upon surgery, 26 were follicular adenomas, 23 follicular carcinomas, and 78 papillary carcinomas. Sections of each sample were evaluated by a pathologist and, when necessary, non-tumoral tissue was dissected. We studied 17 normal thyroid tissues in total, which were obtained from patients with localized disease that underwent hemithyroidectomy and gave consent to take a sample from the unaffected contralateral thyroid lobule. Nine corresponded to tumors included in the study, while for the eight remaining normal thyroid tissues, the matched tumors were not available. At least 80% of the cells were cancerous in all tumor samples, while non-tumor samples had no observable tumor epithelia. Tumor samples were grouped into two sets, the first composed of tumors from *Sant Pau* and *Sabadell* hospitals (16 follicular adenomas, 17 follicular carcinomas, 35 papillary carcinomas, and 8 normal thyroid tissues), and the second composed of tumors from *Arnau de Vilanova* (10 follicular adenomas, 6 follicular carcinomas, 43 papillary tumors, and 9 normal thyroids). The histological classification criteria applied have been previously described.<sup>9</sup>

The clinical follow-up of the patients was carried out by physical examination, neck ultrasonography, simultaneous determination of serum anti-tiroglobulin antibodies with tiroglobulin (basal, or after thyrotropin stimulation by thyroid hormone withdrawal, or the administration of recombinant human thyrotropin), and whole-body iodine scanning. If there was a suspicion of local or distant disease, other imaging techniques such as CT, MRI, PET-CT, or scintigraphy were used. Both structural and biochemical recurrence were considered events in the analysis of recurrence-free survival.

### Genotyping

Using Sanger sequencing, all papillary samples were screened for *BRAF* mutations at codon 600 in exon

15, while follicular adenomas, follicular carcinomas, and follicular variants of papillary tumors were screened for mutations in *H*-, *N*-, and *K-RAS* at mutational hotspots on codons 12 and 13 of exon 2, and codon 61 of exons 3. When available, cDNA from papillary samples was also screened for *RET/PTC1* and *RET/PTC3* rearrangements.

### RNA Extraction and Next-Generation Sequencing

Whole RNA from tumor set 1 was extracted using TRIzol (Life Technologies, MD, USA) according to the manufacturer's instructions. RNAs from tumor set 2 were further purified using RNeasy MinElute Cleanup columns (Qiagen, Valencia, CA, USA). The integrity of RNA was assessed using Agilent BioAnalyzer 2100 (Agilent Technologies).

RNA samples were processed as described in the 'TruSeq Small RNA Sample Preparation Guide' (Illumina part # 15004197 Rev. A of November 2010). Briefly, 1  $\mu$ g of purified total RNA containing the small fraction of RNA was sequentially ligated to 3' and 5' adapters using the truncated form of T4 RNA ligase 2 and the T4 RNA ligase, respectively. Reverse transcription with SuperScript II reverse transcriptase was then used to yield cDNA adapter-ligated libraries that were amplified by PCR with Phusion DNA polymerase and Illumina RNA PCR primers. cDNA-amplified libraries were pooled and separated by polyacrylamide gel electrophoresis, and a fraction of 145–160 bases was extracted. The purified fraction constituted the multiplexed, purified libraries that were applied to an Illumina flow cell to generate clusters, and sequenced on the Genome Analyzer IIX with SBS TruSeq v5 reagents following manufacturer's protocols.

### Statistical Analyses

Detailed statistical methods are presented in Supplementary Materials and Methods. Briefly, after quality assessment,<sup>20</sup> adapter removal,<sup>21</sup> and sequence mapping,<sup>22</sup> differential miRNA expression assessment was carried out using the Bioconductor package edgeR.<sup>23</sup> Benjamini and Hochberg's<sup>24</sup> correction was applied to ensure a false discovery rate (FDR) below 1% and those miRNAs meeting this criterion and with a fold change  $>2$  or  $<0.5$  were considered for further analysis.

For 38 samples of tumor set 1 (see Supplementary Table S1), mRNA expression data were available from previous studies.<sup>10</sup> Taking advantage of the available data from mRNA and miRNA expression from the same tumors, MirRGate, a specific tool for miRNA targets' identification, was used for target prediction (<http://mirgate.bioinfo.cnio.es/API/>). The miRNA was considered regulatory if its expression was negatively correlated with that of the target mRNA as assessed by Pearson's correlation test.

For recurrence-free survival analysis, a shrinking LASSO regression applied via the R package 'glmnet' (v.1.9–8)<sup>25</sup> was used for variable selection. Recurrence-free survival was defined as the time between initial diagnosis and relapse or death due to the disease, with observations censored at the last follow-up if no event had occurred. MiRNA covariates for which the estimated corresponding regression coefficient ( $\beta$ ) did not shrink to zero ( $\beta \neq 0$ ) were considered active coefficients, and evaluated in further steps. An exploratory univariate Cox regression was carried out for these miRNA covariates using the 'survival' package in R (v.2.37–7). Those with FDR  $<0.05$  were considered statistically significant, and evaluated in the multivariate Cox regression model. A signature score (SScore) representative of the signature expression for each sample was calculated as the sum of each miRNA expression level multiplied by the value of the coefficient obtained in the LASSO regression ( $\beta$ ). The median value of the SScore was used to stratify tumors into two groups: high risk and low risk. The final multivariate Cox proportional hazards model was established with a miRNA signature expression based on the goodness of the fit by the Akaike Information Criteria (AIC) along with all remaining covariates, adjusting for the following confounding variables: histological subtype, tumor stage, gender, and age at diagnosis.

## Results

### MiRNA Sequencing Reveals Differential miRNAs Expression Patterns Among Thyroid Tumors

Altogether, data from 127 thyroid tumors including 26 follicular adenomas, 23 follicular carcinomas and 78 papillary carcinomas and 17 normal tissues were generated using next-generation miRNA sequencing. However, since slightly different extraction methods were used for the two tumor sets (see Materials and Methods section), we analyzed them separately, in order to avoid the introduction of technical bias in the results. Tumor set 1 was chosen as discovery series in part because expression profiling data were also available. Tumor set 2 was used to validate results from tumor set 1. The histopathological and mutational characteristics of both tumor sets are summarized in Table 1 (for further information, see Supplementary Table S1).

Unsupervised hierarchical clustering of data from the discovery set revealed the existence of two main clusters (Figure 1a), which resembled those described in a recent DNA methylation study.<sup>9</sup> Follicular tumors (both carcinomas and adenomas) were localized in the same branch, together with normal tissues, while papillary tumors were in a separated branch. Moreover, a separation of samples according to the mutational status was evident. This observation led us to perform a supervised analysis subdividing samples according to both histology and

**Table 1** Summary of the main clinical and pathological characteristics of samples

Clinical characteristics	Tumor set 1 (n = 68) Number (%)	Tumor set 2 (n = 59) Number (%)
<i>Gender<sup>a</sup></i>		
Male	18 (27)	9 (15)
Female	50 (73)	50 (85)
<i>Age<sup>a</sup></i>		
Median	43	49
Min-max	13–77	22–78
<i>Histology<sup>a</sup></i>		
Conventional variant of papillary thyroid carcinoma	26 (38)	31 (53)
Follicular variant of papillary thyroid carcinoma	6 (9)	7 (12)
Other variants of papillary thyroid carcinoma <sup>b</sup>	3 (4)	5 (8)
Follicular thyroid carcinoma	17 (25)	6 (10)
Follicular adenoma	16 (24)	10 (17)
<i>Mutation<sup>a</sup></i>		
<i>BRAF</i> <sup>V600E</sup>	19 (28)	17 (29)
<i>RAS</i>	12 (18)	1 (2)
<i>RET/PTC1</i>	3 (4)	0 (0)
Negative	34 (50)	41 (69)
<i>Recurrence<sup>c</sup></i>		
Yes	16 (31)	14 (29)
No	36 (69)	34 (69)
Missing	0 (0)	1 (2)
<i>Follow-up (months)<sup>d</sup></i>		
Median (interquartile range)	72 (41–96)	24 (10–36)
Normal thyroid tissue	8	9

A total of 127 tumor samples and 17 normal tissues were used in this study, divided in the discovery series (tumor set 1 = 68 tumors and 8 normal thyroid tissues) and replication series (tumor set 2 = 59 tumors and 9 normal samples).

<sup>a</sup>The percentage was calculated taking into account only the total number of tumors (normal tissues were not included).

<sup>b</sup>Other variants of papillary thyroid carcinoma include tall-cell variant, diffuse sclerosing, and oncocytic variant (for more information see Supplementary Table S1).

<sup>c</sup>The data on recurrence are only included for the malignant tumors (neither normal tissues nor adenomas were taken into account).

<sup>d</sup>Median follow-up time and interquartile range were calculated only for disease-free patients with a malignant disease.

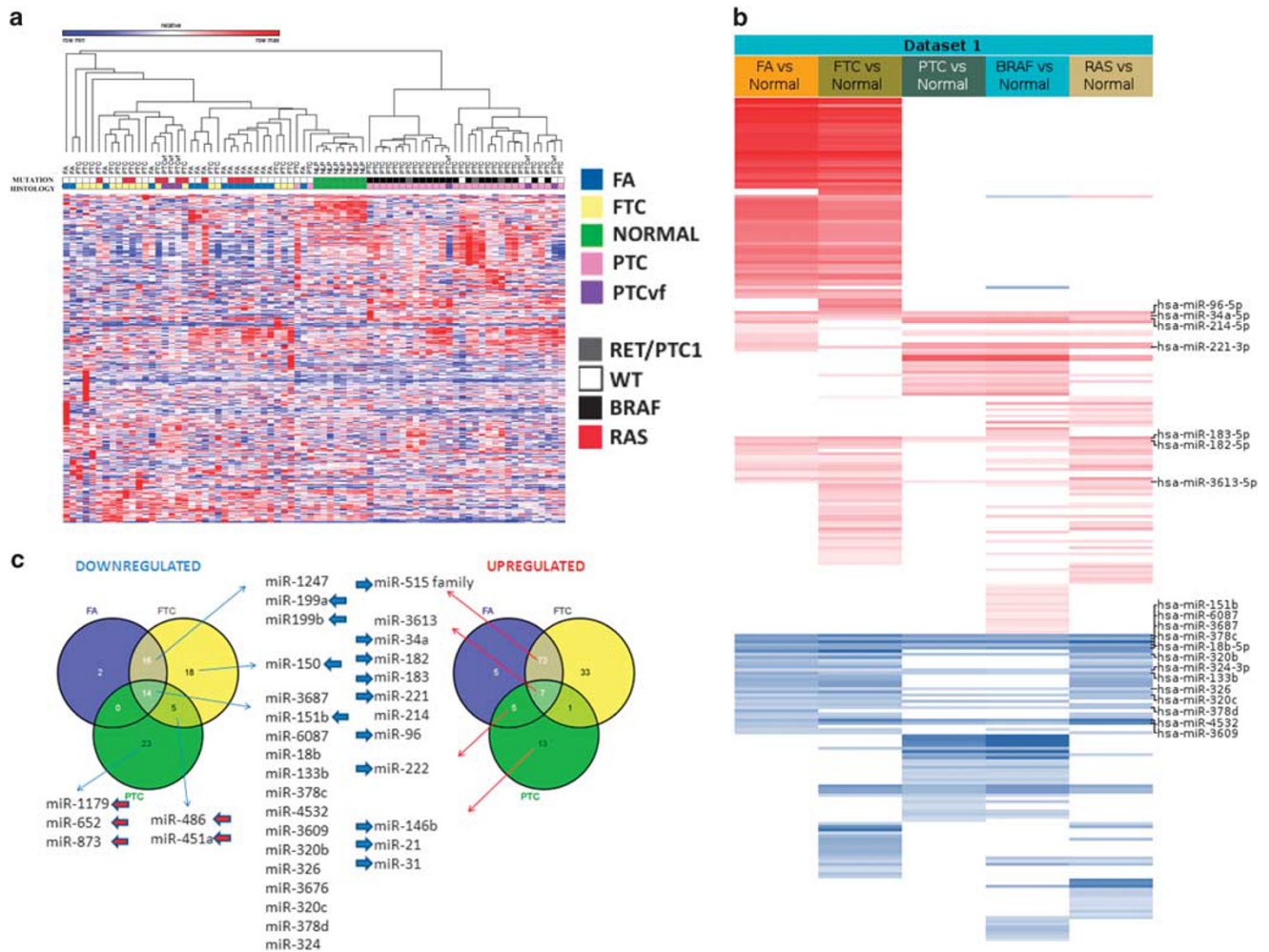
genetics. Of the 808 significantly differentially expressed miRNAs, 170 and 100 were significantly up- and downregulated, respectively. It is worth noting that 7 and 13 miRNAs were consistently up- and downregulated, respectively, in all thyroid tumors (regardless of their malignancy, histological or mutational status; Figure 1b).

The comparison of each histological subtype with normal thyroid samples allowed us to identify a long list of subtype-specifically deregulated miRNAs (Figure 1c). Ninety miRNAs were overexpressed in follicular adenomas by at least twofold compared to normal tissue. There were 114 and 26 overexpressed

in follicular and papillary carcinomas, respectively. The miRNAs with maximum changes exhibited extreme overexpression in follicular adenomas and follicular carcinomas (3452- and 1866-fold, respectively), and 101-fold for papillary samples. Conversely, there were 32 miRNAs downregulated at least twofold in adenomas (maximum 9), 53 in follicular carcinomas (maximum 17), and 42 in papillary tumors (maximum 13). According to a Venn diagram analysis, it seemed papillary samples had a distinct miRNA expression spectrum to that of tumors with a follicular pattern of growth (Figure 1c). Even having relatively few tumors classified as follicular variant of papillary carcinoma, a Venn diagram analysis showed they share more features with follicular tumors than papillary ones (Supplementary Figure S1). Moreover, there was substantial overlap between the deregulated miRNAs identified in follicular adenomas and follicular carcinomas (Figure 1b and c); 89% of upregulated and 94% of downregulated miRNAs identified in the former were also up- and downregulated, respectively, in the latter. A large number of the highly upregulated miRNAs in both of these histogroups belonged to the gene family miR-515 (miR-517a/b, -518a/b/c/e/f, and -516a/b). Other commonly highly upregulated molecules were miR-182, miR-183, and miR-96. The most downregulated miRNA among tumors with a follicular pattern of growth was miR-1247, which has not previously been implicated in thyroid cancer, followed by several members of the miR-199 family (miR-199a/b). Downregulation of miR-150 seemed to be follicular carcinoma specific as it was not detected among adenomas. For samples with a papillary pattern of growth, miR-146b and the miR-221~222 cluster were the most upregulated miRNAs in tumor set 1, followed by miR-21 and -31. MiR-1179 was one of the most downregulated molecules in these tumors. All deregulated miRNAs are listed in Supplementary Table S2.

As the specific driver mutation genotype was available, it was possible to group tumors according to this information, and to compare miRNA expression with that of normal thyroid tissues. Because of low prevalence, *RET/PTC1*-related tumors were not further considered. Figure 1b clearly shows that *RAS*-mutated tumors shared deregulated miRNAs with tumors with a follicular pattern of growth (both adenomas and carcinomas).

This rationale allowed us to identify extended lists of deregulated miRNAs associated with the presence of specific driver alterations. There were 61 upregulated miRNAs in *BRAF*-mutated and 56 in *RAS*-related tumors (showing a maximum 134-fold and 16-fold expression, respectively). However, 53 and 51 miRNAs were downregulated in *BRAF*- and *RAS*-positive samples, respectively (maximum 25- and 17-fold, respectively). In both cases, the most significantly overexpressed miRNAs were those previously described for papillary tumors (miR-146b and the miR-221~222 cluster). Conversely, miR-7 and miR-204 were the most downregulated molecules in



**Figure 1** miRNoma from the discovery series. **(a)** Unsupervised hierarchical cluster analysis of 68 primary thyroid tumors and 8 normal tissue samples divided the sample set into two main clusters. ‘Branch 1’ was composed of all follicular carcinomas (FTC) and follicular adenomas (FA) as well as all normal tissues. ‘Branch 2’ included the majority of papillary samples (PTC). Separation of samples according to the driver mutation was also apparent. **(b)** Heatmap representation of differentially expressed miRNAs. Tumors were compared to normal thyroid tissues based on histological as well as driver mutation grouping, which was apparent from the unsupervised analysis. Color legend: red, upregulated; blue, downregulated. The 13 down- and 7 up- commonly deregulated miRNAs among all tumor classes (regardless of the malignancy and mutational status) are listed. **(c)** Venn diagram analysis of subtype-specific differentially expressed miRNAs revealed substantial overlap between follicular adenomas (FA) and follicular carcinomas (FTC), while papillary samples (PTC) showed a distinct miRNA expression signature. Moreover, 14 miRNAs were commonly upregulated while 7 were downregulated among all three subtypes. Only miR-3676 was not present when the mutation was considered **(b)**. Blue arrows denote miRNAs previously described elsewhere to be implicated in thyroid cancer. Red arrows denote molecules described in Swierniak *et al.*<sup>16</sup>

*BRAF* tumors; the latter was statistically significant when *BRAF* tumors were compared with wild-type papillary samples (data not shown). All deregulated miRNAs are listed in Supplementary Table S2.

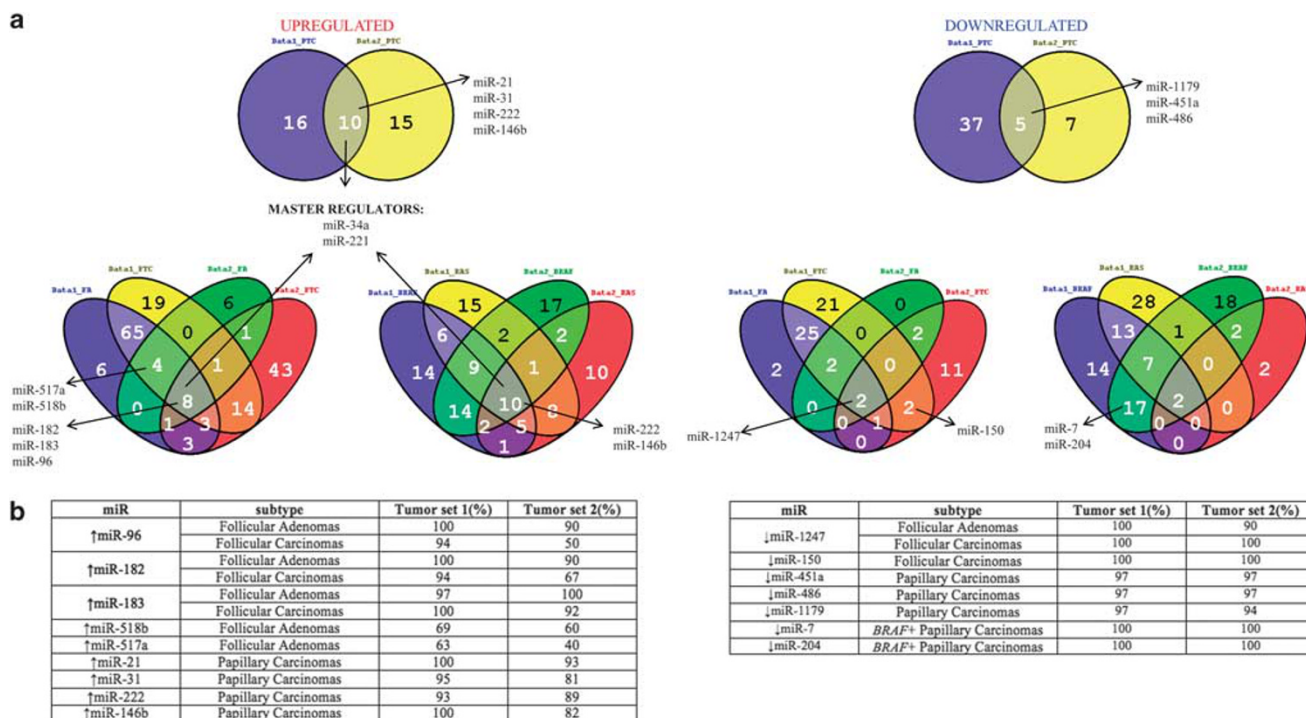
### Validation of Deregulated miRNAs

A smaller number of deregulated miRNAs was observed in tumor set 2, possibly due to the slightly different RNA extraction procedure followed. Nevertheless, two out of the three commonly upregulated miRNAs had been already detected in tumor set 1 (miR-221 and miR-34a), termed from now on ‘master regulators’.

When it comes to the subtype-specifically deregulated miRNAs, a considerable proportion of these had

been already detected in the discovery phase (62, 34, and 43% among follicular adenoma-, follicular thyroid carcinoma-, and papillary thyroid carcinoma-specific probes; Figure 2a and Supplementary Table S3). It seemed that the grouping of samples according to the mutational status was more robust, as a higher proportion of miRNAs was validating the results of tumor set 1 (60 and 58% in *BRAF* and *RAS*, respectively; Figure 2a and Supplementary Table S3). Commonly deregulated miRNAs in both tumor sets are summarized in Supplementary Table S4.

As detailed in the Figure 2, upregulation of miR-96, -182, and -183 was confirmed for tumors with a follicular pattern of growth, while a considerable increase in some members of the miR-515 family was only detected among follicular adenomas



**Figure 2** Validation step. (a) Venn diagram analysis identified those molecules commonly detected in both tumor sets and therefore considered validated. (b) Proportion of tumors of both sets showing deregulation of some of the commonly deregulated miRNAs. Arrow next to the miRNA name indicates whether the molecule was detected as downregulated (↓) or upregulated (↑).

of tumor set 2. Significant downregulation of miR-1247 was observed for tumors with follicular pattern of growth in both tumor sets, while, miR-150 downregulation was validated to be follicular carcinoma specific. In papillary tumors, several upregulated (miR-21, -31, -222, and -146b) as well as downregulated (miR-451a, -486, and -1179) molecules were validated in the tumor set 2. The downregulation of miR-7 and -204 was again strongly associated with the presence of *BRAF* mutation in the validation set. The proportions of tumors showing these deregulations are detailed in Figure 2b.

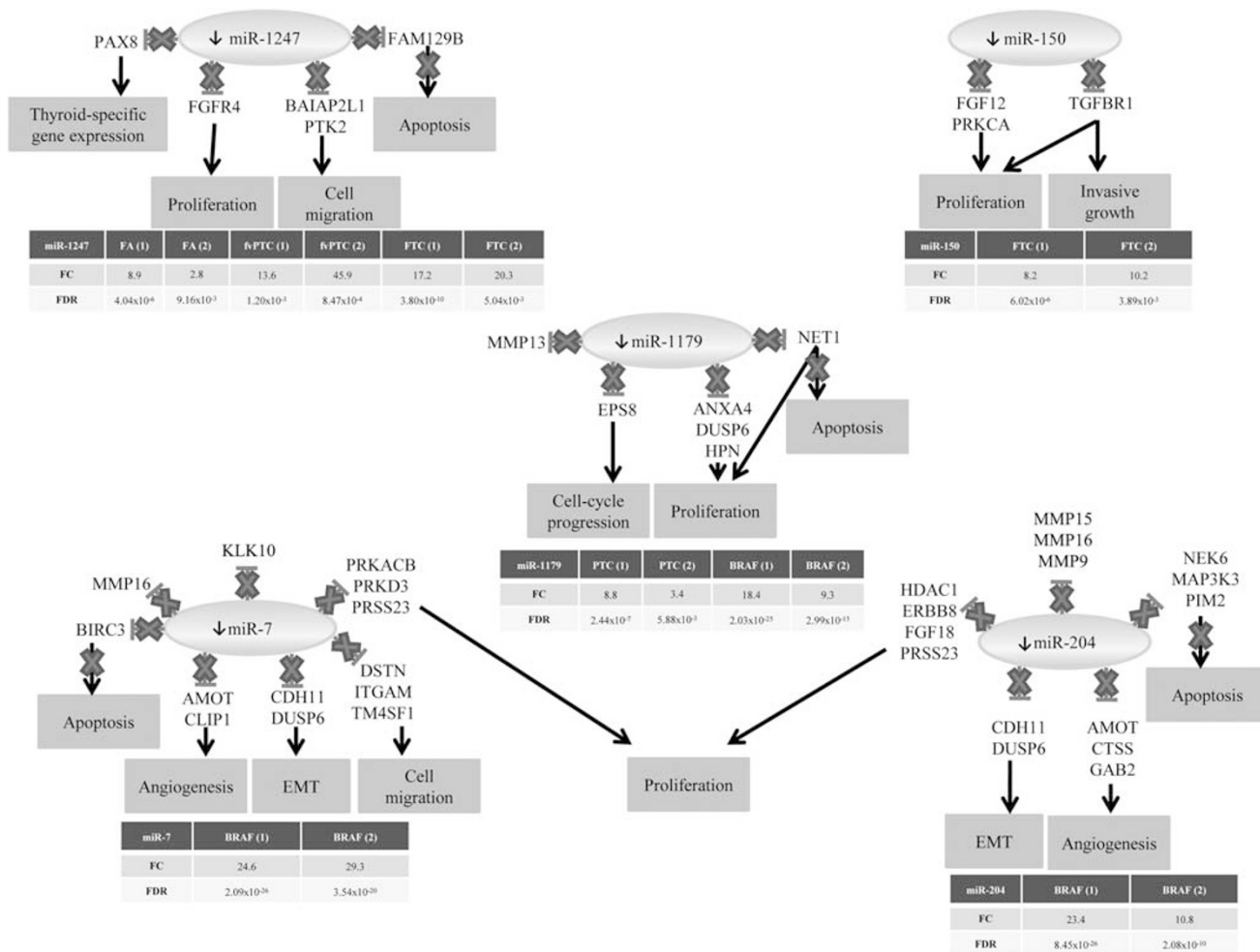
**Integrated Gene Expression Reveals Possible Targets for Novel miRNAs**

For those novel deregulated miRNAs common to both tumor sets, we integrated the miRNAseq data with gene expression data available for the tumors from tumor set 1 using miRGate software (Figure 3). MiR-1247 was less expressed in all tumors with a follicular pattern of growth (follicular adenomas, follicular carcinomas, and follicular variant of papillary tumors) than in normal tissue (fold change 3–46). According to validated and predicted interaction data available, the negative regulation of relevant targets could be impaired. These include molecules involved in cell proliferation (*FGFR4*), migration (*BAIAP2L1* and *PTK2*), apoptosis (*FAM129B*), as well as a thyroid-specific transcription factor

(*PAX8*). MiR-150 was downregulated in follicular thyroid carcinomas (8- to 10-fold change) and its expression correlated negatively with genes involved in proliferation (*FGF12*, *PRKCA*, and *TGFBR1*) and invasive growth (*TGFBR1*). The compromised expression of miR-1179 seen in papillary tumors (three- to ninefold change) could lead to the deregulation of molecules involved in cell cycle progression (*ESP8* and *ANXA4*) or invasiveness (*HPN* and *MMP13*). Finally, miR-7 and miR-204 were severely downregulated in *BRAF*-positive samples (25- to 29- and 11- to 23-fold change, respectively), potentially affecting a long list of possible targets, including genes involved in extracellular matrix remodeling (*KLK10*, *MMP9*, *MMP15*, and *MMP16*), angiogenesis (*AMOT*, *CLIP1*, *CTSS*, and *GAB2*), and epithelial-mesenchymal transition (*CDH11* and *DUSP6*).

**A 2-Genes miRNA Signature Associated with Recurrence**

Using the linear regression selection method LASSO, 16 miRNAs were selected for inclusion in the recurrence-free survival analysis. Univariate analysis and correction for multiple testing pointed to the relevance of five of them (let-7a-3p, miR-30e-5p, miR-192-3p, let-7d-3p, and miR-493-5p). Applying Akaike Information Criteria the best model included two molecules (miR-192 and let-7a) whose expression individually (Figure 4a), but even more



**Figure 3** Predicted and validated targets of novel miRNAs. MiRGate software was applied to identify possible targets of novel molecules identified in this study. A small downward arrow to the left of specific microRNAs refers to downregulation. Larger arrows denote where downregulation of the miRNA gives rise to a lack of inhibition of the target gene pointed to. (1) refers to results from tumor set 1, while (2) refers to those from tumor set 2. EMT, epithelial-mesenchymal transition; FA, follicular adenoma; FC, fold change; FDR, false discovery rate; FTC, follicular thyroid carcinoma; fvPTC, follicular variant of papillary thyroid carcinoma; PTC, papillary thyroid carcinoma.

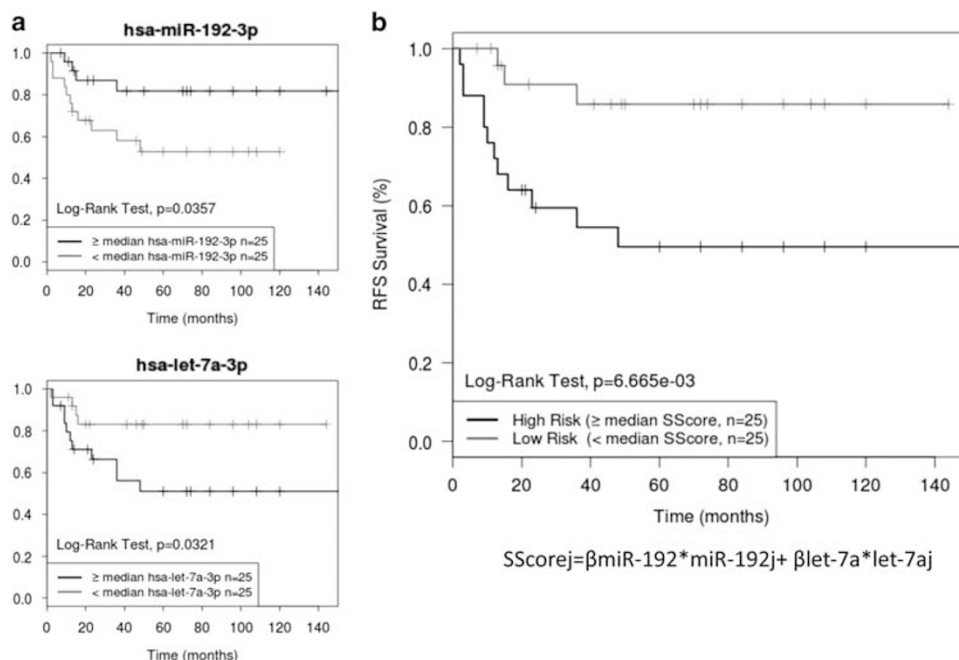
significantly in combination (Figure 4b), discriminated patients with a higher probability of recurrence. In this model, a combination of increased expression of let-7a, together with decreased miR-192 expression, was associated with an increased risk of recurrence. Even after correcting for other important clinical features, such as subtype, stage, gender, and age, the miRNAs remained significant predictors of disease relapse in the final prediction model (Supplementary Figure S2).

### Discussion

The currently available evidence suggests that miRNA deregulation could be a crucial event in thyroid carcinogenesis. Most of what is known about deregulated miRNAs has been based on microarray profiling or real-time PCR, both of which have inherent biases. Approaches based on next-

generation sequencing permit the detection and simultaneous quantification of the miRNAs present in a specimen,<sup>26</sup> thus affording an in-depth, unbiased characterization of full miRNomes. In the present work, using two independent series of well-differentiated thyroid tumors and next-generation sequencing, we were able to identify deregulated miRNAs related to the specific histological subtypes and driver mutations. These included not only miRNAs previously reported to be related to thyroid cancer, but also novel recurrent markers. We propose possible targets for these novel molecules based on *in silico* prediction using miRNA and mRNA expression data from the same tumors. Finally, we describe a model based on a two-miRNA signature that could be used in relapse prediction.

Two miRNA molecules, miR-221 and miR-34a, were found to be consistently deregulated among all the tumor groups studied, suggesting that they play a key role, both in the first steps of the malignant



**Figure 4** Markers of recurrence-free survival. (a) Individual Kaplan–Meier curves for each signature member (miR-192 and let-7a) with expression dichotomized at the median value. (b) Kaplan–Meier survival analysis. A signature score (SScore) representing the signature expression for each sample was calculated as  $SScore_j = \beta_{miR-192} \times miR-192_j + \beta_{let-7a} \times let-7a_j$ , where  $\beta_x$  is the coefficient for  $x$  obtained from the LASSO regression. The score was dichotomized at the median to create two groups for comparison.

transformation process in thyroid carcinogenesis, and in more advanced forms of the disease. That is, they may act as master regulators in thyroid cancer. The involvement of miR-221 in carcinogenesis has already been studied, as its expression is increased in various neoplasias. Our results confirm those of a previous microarray study, where the authors observed upregulation of miR-221 in various histological subtypes of thyroid cancer (follicular adenomas, follicular carcinomas, and papillary carcinomas).<sup>14</sup> In thyroid cells, it was shown that miR-221 has oncogenic properties, as it induces cell cycle progression by targeting p27<sup>Kip1</sup>.<sup>27</sup> To date, the overexpression of miR-34a has been described in both papillary tumors<sup>12,17</sup> and papillary thyroid carcinoma cell lines.<sup>28</sup> Recently, it has been suggested that its oncogenic effect in PTC involves activation of the PI3K/Akt/Bad pathway,<sup>29</sup> which is consistent with it having a master regulator role. Moreover, both miR-34a and -221 have been very recently identified in the TCGA data set as crucial regulators of the immune response activities among all papillary carcinoma variants studied by this consortium,<sup>19</sup> further confirming our results.

As observed with other genomic features,<sup>4,9</sup> the miRNomes of tumors with a follicular pattern of growth were largely similar, pointing to a possible progressive evolution of carcinomas from follicular adenomas.<sup>30</sup> The observed upregulation of miR-96, miR-182, and miR-183 in both follicular adenomas and follicular carcinomas has been reported previously.<sup>14,31,32</sup> Of the other molecules that had been previously reported to be deregulated in thyroid

cancer,<sup>32</sup> it is noteworthy that several members of the miR-515 family were highly overexpressed in both follicular adenomas (showing a maximum fold change of 3452) and follicular carcinomas (maximum fold change of 1867) in tumor set 1. However, for few of those was that overexpression replicated in tumor set 2. Similarly, downregulation of the miR-199 family was detected in the discovery phase of our study, confirming previously published findings, but not in the validation phase.<sup>31,32</sup> It is likely that the use of a purification column introduces bias into miRNA profiles. Indeed, the effect of extraction technique on miRNA expression detection has been previously reported by others.<sup>33</sup> The identification of deregulated miRNAs common to both tumor sets, despite the different methodologies employed, adds weight to the evidence that they play a role in thyroid cancer.

One interesting novel marker was miR-1247, which was found to be downregulated in all tumors with a follicular pattern of growth (including even adenomas), indicating it could be an early event in the development of these neoplasias. The lack of negative regulation of its predicted targets could give proliferative potential to the follicular cell. Indeed, it has recently been confirmed that this molecule plays a tumor-suppressive role in pancreatic cancer by inhibiting proliferation and tumorigenesis and triggering G<sub>0</sub>/G<sub>1</sub> cell cycle arrest.<sup>34</sup> Further, miR-150 was downregulated specifically in follicular carcinomas. In line with its recently described downregulation in poorly differentiated thyroid tumors,<sup>35</sup> the candidate targets identified here (*FGF12*, *PRKCA*,



and *TGFBR1*) could lead to a more aggressive neoplasm.

As reported in tumors with follicular growth pattern, we observed upregulation in papillary tumors of previously described markers, such as miR-146b, the miR-221~222 cluster, miR-21 and miR-31.<sup>12,14,15,17</sup> Several other miRNAs (miR-486, -873, -1179, -451a, and -652), firstly described in the first next-generation sequencing study performed in papillary thyroid carcinoma,<sup>16</sup> were also detected. This consistency with findings from earlier studies served as an external validation of our results. Particularly noteworthy was the consistent miR-204 downregulation observed in both tumor sets, specifically associated with the presence of the *BRAF* mutation. The targets of this miRNA were predicted to impair several cellular processes, such as those maintaining epithelial physiology,<sup>36</sup> which could explain the relatively poor prognosis described for these patients.<sup>37</sup> Another interesting marker was miR-7, highly significantly associated with *BRAF* mutation when compared to normal thyroid tissues in tumor set 1, and with wild-type papillary carcinomas in the validation set. Moreover, one of the predicted targets of miR-7 was *KLK10*, which our group recently identified as specifically hypomethylated in *BRAF*-mutated tumors and correlated with gene expression.<sup>9</sup> Thus, overexpression of miR-7 represents a second mechanism by which this molecule may be specifically upregulated in *BRAF*-positive papillary thyroid tumors, thereby highlighting its importance.

Though most well-differentiated tumors can be effectively clinically managed, there is a subset of patients who develop recurrences and often respond poorly to the current therapeutic options. Here, we propose a relapse prediction model based on a two-miRNA signature composed of a combination of let-7a increase together with a miR-192 decrease. The let-7 family of miRNAs are well-established tumor suppressors;<sup>38</sup> their increased expression has been detected in the serum from papillary thyroid cancer patients (compared to healthy individuals or benign cases) as associated with the presence of multifocal lesions ( $P < 0.001$ ).<sup>39</sup> Thus, further studies are required to elucidate the role of this miRNA family in thyroid tumorigenesis. However, the tumor suppressive role of miR-192 has been robustly functionally demonstrated,<sup>40–42</sup> and is consistent with our current observation. In this regard, the value of the tumor set 2 was limited for validation, as the time of follow-up was shorter. Thus, while the predictive value of our signature requires validation, it seems that these miRNAs could be considered as markers of recurrence, acting independently of other known prognostic factors, and therefore potentially contributing to an improved stratification of thyroid cancer patients.

To summarize, this is the first next-generation sequencing study to be applied to all major histological subtypes of well-differentiated thyroid cancer in the largest tumor collection reported so far. We

update the current knowledge about deregulated miRNAs by confirming some of those that have been previously reported as well as by describing novel ones. According to our results, there are at least two miRNA molecules (miR-34a and miR-221) acting as master regulators of thyroid carcinogenesis. Moreover, by integrating miRNA and mRNA expression data for the same tumors, we were able to explore the possible targets of the novel molecules detected, and thus shed further light on the biological mechanisms involved in thyroid carcinogenesis. By integrating clinicopathological data with the miRNA expression, we identified a two-miRNA signature associated with disease recurrence.

## Acknowledgments

This work was supported in part by the Fondo de Investigaciones Sanitarias (projects PI11/01359, PI14/00240, PI11/01354, RD12/0036/0030 and RD12/0036/0013), the Fundacion Mutua Madrilenia (project AP2775/2008), the Comunidad de Madrid S2011/BMD-2328 TIRONET, and the Spanish Ministry of Economy and Competitiveness (projects SAF2011/23638 and SAF2013/44709R). VM and AAdC are predoctoral fellows of 'la Caixa'/CNIO international PhD programme. LI-P is a predoctoral fellow of the CIBERER.

## Disclosure/conflict of interest

The authors declare no conflict of interest.

## References

- 1 Calin GA, Croce CM. MicroRNA signatures in human cancers. *Nat Rev Cancer* 2006;6:857–866.
- 2 Calin GA, Ferracin M, Cimmino A *et al*. A microRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N Engl J Med* 2005;353:1793–1801.
- 3 Lu J, Getz G, Miska EA *et al*. MicroRNA expression profiles classify human cancers. *Nature* 2005;435:834–838.
- 4 Rossing M. Classification of follicular cell-derived thyroid cancer by global RNA profiling. *J Mol Endocrinol* 2013;50:R39–R51.
- 5 Nikiforov YE, Nikiforova MN. Molecular genetics and diagnosis of thyroid cancer. *Nat Rev Endocrinol* 2011;7:569–580.
- 6 Kebebew E, Greenspan FS, Clark OH *et al*. Anaplastic thyroid carcinoma. Treatment outcome and prognostic factors. *Cancer* 2005;103:1330–1335.
- 7 DeLellis R, Lloyd R, Heitz P *et al*. Pathology and genetics of tumours of endocrine organs. IARC Press: Lyon. 2004.
- 8 Giordano TJ, Kuick R, Thomas DG *et al*. Molecular classification of papillary thyroid carcinoma: distinct *BRAF*, *RAS*, and *RET/PTC* mutation-specific gene expression profiles discovered by DNA microarray analysis. *Oncogene* 2005;24:6646–6656.

- 9 Mancikova V, Buj R, Castelblanco E *et al*. DNA methylation profiling of well-differentiated thyroid cancer uncovers markers of recurrence free survival. *Int J Cancer* 2014;135:598–610.
- 10 Montero-Conde C, Martin-Campos JM, Lerma E *et al*. Molecular profiling related to poor prognosis in thyroid carcinoma. Combining gene expression data and biological information. *Oncogene* 2008;27:1554–1561.
- 11 Pallante P, Battista S, Pierantoni GM *et al*. Deregulation of microRNA expression in thyroid neoplasias. *Nat Rev Endocrinol* 2014;10:88–101.
- 12 He H, Jazdzewski K, Li W *et al*. The role of microRNA genes in papillary thyroid carcinoma. *Proc Natl Acad Sci USA* 2005;102:19075–19080.
- 13 Jazdzewski K, Murray EL, Franssila K *et al*. Common SNP in pre-miR-146a decreases mature miR expression and predisposes to papillary thyroid carcinoma. *Proc Natl Acad Sci USA* 2008;105:7269–7274.
- 14 Nikiforova MN, Tseng GC, Steward D *et al*. MicroRNA expression profiling of thyroid tumors: biological significance and diagnostic utility. *J Clin Endocrinol Metab* 2008;93:1600–1608.
- 15 Pallante P, Visone R, Ferracin M *et al*. MicroRNA deregulation in human thyroid papillary carcinomas. *Endocr Relat Cancer* 2006;13:497–508.
- 16 Swierniak M, Wojcicka A, Czetwertynska M *et al*. In-depth characterization of the microRNA transcriptome in normal thyroid and papillary thyroid carcinoma. *J Clin Endocrinol Metab* 2013;98:E1401–E1409.
- 17 Tetzlaff MT, Liu A, Xu X *et al*. Differential expression of miRNAs in papillary thyroid carcinoma compared to multinodular goiter using formalin fixed paraffin embedded tissues. *Endocr Pathol* 2007;18:163–173.
- 18 Cancer Genome Atlas Research Network Integrated genomic characterization of papillary thyroid carcinoma. *Cell* 2014;159:676–690.
- 19 Huang CT, Oyang YJ, Huang HC *et al*. MicroRNA-mediated networks underlie immune response regulation in papillary thyroid carcinoma. *Sci Rep* 2014;4:6495.
- 20 Andrews S. FastQC a quality-control tool for high-throughput sequence data <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- 21 Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 2011;17:10–12.
- 22 Langmead B, Trapnell C, Pop M *et al*. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009;10:R25.
- 23 Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010;26:139–140.
- 24 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 1995;57:289–300.
- 25 Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 2010;33:1–22.
- 26 Morozova O, Marra MA. Applications of next-generation sequencing technologies in functional genomics. *Genomics* 2008;92:255–264.
- 27 Visone R, Russo L, Pallante P *et al*. MicroRNAs (miR)-221 and miR-222, both overexpressed in human thyroid papillary carcinomas, regulate p27Kip1 protein levels and cell cycle. *Endocr Relat Cancer* 2007;14:791–798.
- 28 Cahill S, Smyth P, Finn SP *et al*. Effect of ret/PTC 1 rearrangement on transcription and post-transcriptional regulation in a papillary thyroid carcinoma model. *Mol Cancer* 2006;5:70.
- 29 Ma Y, Qin H, Cui Y. MiR-34a targets GAS1 to promote cell proliferation and inhibit apoptosis in papillary thyroid carcinoma via PI3K/Akt/Bad pathway. *Biochem Biophys Res Commun* 2013;441:958–963.
- 30 Arora N, Scognamiglio T, Zhu B *et al*. Do benign thyroid nodules have malignant potential? An evidence-based review. *World J Surg* 2008;32:1237–1246.
- 31 Dettmer M, Vogetseder A, Durso MB *et al*. MicroRNA expression array identifies novel diagnostic markers for conventional and oncocytic follicular thyroid carcinomas. *J Clin Endocrinol Metab* 2013;98:E1–E7.
- 32 Rossing M, Borup R, Henao R *et al*. Down-regulation of microRNAs controlling tumourigenic factors in follicular thyroid carcinoma. *J Mol Endocrinol* 2012;48:11–23.
- 33 Kim YK, Yeo J, Kim B *et al*. Short structured RNAs with low GC content are selectively lost during extraction from a small number of cells. *Mol Cell* 2012;46:893–895.
- 34 Shi S, Lu Y, Qin Y *et al*. miR-1247 is correlated with prognosis of pancreatic cancer and inhibits cell proliferation by targeting neuropilins. *Curr Mol Med* 2014;14:316–327.
- 35 Dettmer MS, Perren A, Moch H *et al*. MicroRNA profile of poorly differentiated thyroid carcinomas: new diagnostic and prognostic insights. *J Mol Endocrinol* 2014;52:181–189.
- 36 Wang FE, Zhang C, Maminishkis A *et al*. MicroRNA-204/211 alters epithelial physiology. *FASEB J* 2010;24:1552–1571.
- 37 Xing M, Westra WH, Tufano RP *et al*. BRAF mutation predicts a poorer clinical prognosis for papillary thyroid cancer. *J Clin Endocrinol Metab* 2005;90:6373–6379.
- 38 Johnson SM, Grosshans H, Shingara J *et al*. RAS is regulated by the let-7 microRNA family. *Cell* 2005;120:635–647.
- 39 Yu S, Liu Y, Wang J *et al*. Circulating microRNA profiles as potential biomarkers for diagnosis of papillary thyroid carcinoma. *J Clin Endocrinol Metab* 2012;97:2084–2092.
- 40 Feng S, Cong S, Zhang X *et al*. MicroRNA-192 targeting retinoblastoma 1 inhibits cell proliferation and induces cell apoptosis in lung cancer cells. *Nucleic Acids Res* 2011;39:6669–6678.
- 41 Kim T, Veronese A, Pichiorri F *et al*. p53 regulates epithelial-mesenchymal transition through microRNAs targeting ZEB1 and ZEB2. *J Exp Med* 2011;208:875–883.
- 42 Song B, Wang Y, Kudo K *et al*. miR-192 Regulates dihydrofolate reductase and cellular proliferation through the p53-microRNA circuit. *Clin Cancer Res* 2008;14:8080–8086.

Supplementary Information accompanies the paper on Modern Pathology website (<http://www.nature.com/modpathol>)