

RESEARCH

Open Access

# Validity of multiplex biomarker model of 6 genes for the differential diagnosis of thyroid nodules

Kristine Ducena<sup>1\*</sup>, Arturs Ābols<sup>1,2</sup>, Janis Vilmanis<sup>3</sup>, Zenons Narbutis<sup>3</sup>, Juris Tārs<sup>4</sup>, Diana Andrējeva<sup>2</sup>, Aija Linē<sup>2</sup> and Valdis Pīrāgs<sup>1,3\*</sup>

## Abstract

**Background:** Currently the cytological examination of fine needle aspiration (FNA) biopsies is the standard technique for the pre-operative differential diagnosis of thyroid nodules. However, the results may be non-informative in ~20% of cases due to an inadequate sampling and the lack of highly specific, measurable cytological criteria, therefore ancillary biomarkers that could aid in these cases are clearly needed. The aim of our study was to evaluate the mRNA expression levels of 8 candidate marker genes as the diagnostic biomarkers for the discrimination of benign and malignant thyroid nodules and to find a combination of biomarkers with the highest diagnostic value.

**Materials and methods:** mRNA expression levels of eight candidate marker genes - *BIRC5*, *CCND1*, *CDH1*, *CITED1*, *DPP4*, *LGALS3*, *MET* and *TFF3* was measured by real-time RT-PCR in paired nodular and surrounding normal thyroid tissue specimens of 105 consecutive patients undergoing thyroid surgery and compared between different types of thyroid lesions.

**Results:** Significant differences in the mRNA expression levels between the normal and malignant thyroid tissues and between benign and malignant nodules were found for *BIRC5*, *CCND1*, *CITED1*, *DPP4*, *LGALS3*, *MET* and *TFF3*, but not *CDH1*. On a single gene basis, relative quantity (RQ) of *LGALS3* had the highest diagnostic value for the discrimination of malignant and benign thyroid nodules (AUC = 0.832,  $P < 0.0001$  and 90.9% sensitivity and 65.6% specificity at the optimal cut-off on ROC curve). The only two-marker set that outperformed *LGALS3* was RQ sum of *LGALS3* and *BIRC5* (AUC = 0.841,  $P < 0.0001$ ). An application of multivariate logistic regression analysis resulted in the generation of a multiplex biomarker model based on *LGALS3*, *BIRC5*, *TFF3*, *CCND1*, *MET* and *CITED1* that had considerably higher specificity than a single marker or two marker gene-based models (AUC = 0.895,  $P < 0.0001$ , 70.5% sensitivity and 93.4% specificity).

**Conclusions:** This study confirmed that mRNA expression levels of 7 out of 8 candidate genes analysed have a diagnostic value for the distinction of benign and malignant thyroid nodules. The multiplex biomarker model based on 6 genes outperformed a single marker or two marker-based models and warrants feasibility studies on FNA biopsies and the validation in a larger cohort of patients.

## Background

Thyroid nodules are a very common clinical finding - the prevalence of palpable nodules ranges from 3 to 7% in the general population but can be as high as 50% based on ultrasonography or autopsy data [1]. Although only less than 5% of the palpable nodules are malignant lesions, thyroid cancers are the most common

malignancy of endocrine organs. According to European Cancer Observatory data age-standardised incidence rate was 3.1 cases and 8.8 cases per 100 000 in men and women, respectively in Europe in 2008 and its incidence rates have steadily increased over the recent decades [2,3]. More than 95% of malignant lesions are derived from thyroid follicular cells and are divided into papillary and follicular carcinomas that differ mainly in the mode of metastatic spread (lymphatic and haematogenous spread, respectively) yet both are relatively indolent tumours with 5-year survival rates > 90%, and

\* Correspondence: [kducena@inbox.lv](mailto:kducena@inbox.lv); [pirags@latnet.lv](mailto:pirags@latnet.lv)

<sup>1</sup>Faculty of Medicine, University of Latvia, Raina Bulvaris 19, Riga, LV1586, Latvia

Full list of author information is available at the end of the article

undifferentiated or anaplastic carcinoma that is a highly aggressive and lethal cancer with 5-year survival rate below 1-17% [4]. A minority of thyroid carcinomas are derived from parafollicular cells (or C-cells) and are referred to as medullary carcinomas with 5-year survival rate of ~80% [4]. Benign nodules include hyperplastic follicular adenomas, multinodular goiter, thyroiditis and benign cysts [1].

The differential diagnosis of nodular thyroid disease is based on cytological examination of fine needle aspiration (FNA) biopsies, however the results may be non-informative in ~20% of cases due to an inadequate sampling and the lack of highly specific, measurable cytological criteria [4,5]. Furthermore, the principal diagnostic feature for follicular thyroid carcinomas (FTC) is capsular or vascular invasion and therefore it can not be reliably distinguished from follicular adenomas by analysis of FNA smears and the definite diagnosis relies on the histological examination of the postoperative surgical specimens [6]. Hence molecular biomarkers of malignancy that could reliably discriminate between malignant and benign nodules in the grey zone of thyroid FNA and classify tumours into the histological subtypes are clearly needed.

Gene expression profiling using cDNA microarrays or serial analysis of gene expression (SAGE) has revealed several hundreds of genes that are differentially expressed between malignant and benign thyroid nodules [7-9]. A number of them, including *LGALS3*, *MET* etc have been shown to be functionally involved in the carcinogenic process, have been validated by qRT-PCR and confirmed at protein level by immunohistochemistry [10-12]. However, several studies have demonstrated that none of these genes individually has sufficient sensitivity and specificity to be exploited as an independent diagnostic biomarker [13,14] while using very large panels of genes may not be practical. Therefore the definition of minimal set of marker genes that allows the correct classification of nodules is required in order to develop a clinically applicable biomarker assay. In the current study, we evaluated the diagnostic value of 8 candidate marker genes - *BIRC5*, *CCND1*, *CDH1*, *CITED1*, *DPP4*, *LGALS3*, *MET* and *TFF3* by analysing their mRNA expression levels in nodular and adjacent relatively normal thyroid tissue specimens of 105 consecutive patients with thyroid nodules who underwent thyroidectomy, and developed a multiplex biomarker model for the classification of the nodules.

## Methods

### Patients and tissue specimens

Paired specimens of thyroid nodules and surrounding normal tissues were collected from 105 consecutive patients undergoing total or partial thyroidectomy at

the Latvian Centre of Oncology and Pauls Stradiņš Clinical University Hospital during the period 2009-2010. Tissue samples were macroscopically dissected by histopathologist during the surgery and stored in RNALater® (Applied Biosystems, USA) at -20°C until processing. Tissue sections were evaluated by experienced pathologist and the diagnosis was established according to standard histopathological criteria. Sixty one of the nodules were diagnosed as follicular adenomas (FA), 33 as papillary thyroid carcinomas (PTC), 5 as medullary thyroid carcinomas (MTC), 3 as anaplastic thyroid carcinomas (ATC) and 3 as follicular thyroid carcinomas (FTC). Selected clinical-pathological data of the patient groups are provided in Table 1. The specimens were collected after the patients' informed consent was obtained in accordance with the regulations of Ethics Committee of the University of Latvia Institute of Experimental and Clinical Medicine.

**Table 1 Clinical and pathological characteristics of the patient groups**

Characteristics	Thyroid cancer n = 44	Benign nodule n = 61
<b>Gender</b>		
Male	9	4
Female	35	57
<b>Age at diagnosis</b>		
Mean ± SD	55 ± 17	54 ± 15
Median (range)	55 (24-83)	57 (25-83)
<b>Family history of cancer</b>		
	37	49
<b>TSH level µIU/mL</b>		
Mean ± SD	2,31 ± 1,41	0,87 ± 0,74
Median (range)	1,32 (1,1-4,34)	0,75 (0,004-2,43)
<b>US criteria</b>		
Hypoechogetic nodule	9	14
No Halozones	0	1
Irregular frontier (non-homogenic)	6	14
Microcalcينات	4	9
Nodules > 3 cm	5	19
Central vascularization	3	3
Retrosternal	1	7
Swollen lymph nodes	2	0
No data	18	8
<b>Histopathology</b>		
PTC	33	
FTC	3	
MTC	5	
ATC	3	
FA		61

### RNA extraction and cDNA synthesis

Tissue specimens were homogenised using the FastPrep-24 instrument and Lysing Matrix D (MP Biomedicals, USA) in 1 ml of Lysis solution<sup>®</sup> (Ambion, USA) followed by the extraction of total RNA using MirVana<sup>®</sup> (Ambion, USA) according to manufacturer's instructions. RNA extracted from tissue material was treated with DNase prior to cDNA synthesis (Ambion, USA) and quantified by NanoDrop ND-100 spectrophotometer. cDNA was synthesized by random hexamer priming from 4 µg of total RNA by using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, USA) according to manufacturer's instructions.

### Quantitative RT-PCR

Quantitative RT-PCR (qPCR) reactions were performed using 2 µl of 1:10 diluted cDNA reaction mixtures, Absolute Blue<sup>™</sup> SYBR green Low ROX (Thermo Scientific, USA) on ABI7500 sequence detection system (Applied Biosystems, USA). Appropriate primer concentrations were established by cDNA 4 log serial dilution curves to ensure amplification efficiency over 95% and the specificity of the amplification products were verified by the melting curve analysis. Sequences of primers used in this study are provided in Table 2. qPCR conditions were as follows: hold for 15 minutes at 95°C, follow by two step PCR for

40 cycles of 95°C for 15 seconds, followed by 60°C for 1 minute. All reactions were performed in duplicates. To choose the appropriate reference genes, a set of seven candidate reference genes (*PGK1*, *PLA2*, *TUBA3*, *ACTB*, *GAPDH*, *TBP*, *POLR2A*) was tested for their mRNA expression stability in 10 thyroid tissue samples from 5 patients (2 PTC, 3 FA, 5 normal) and the two most stable genes (*POLR2A* and *PGK1*) were determined by using the open-access software GeNorm Win3.4 (<http://medgen.ugent.be/~jvdesomp/genorm>). Normalisation factor (NF) for each cDNA sample was then calculated by using GeNorm Win3.4 based on *PGK1* and *POLR2A* expression level. Pooled cDNA from 10 thyroid tissue samples was used as a reference sample and included in each experiment to allow the comparison across multiple plates. The expression level of each gene was determined relative to its expression in the sample with the highest expression (lowest Ct value). The analysis of relative gene expression (RQ) data was performed using the  $2^{-\Delta\text{CT}}$  method and to eliminate variations of cDNA quantity and quality the data were normalised by using NF for each sample ( $\text{RQ} = (2^{-\Delta\text{CT}})/\text{NF}$ ).

### Statistical analysis

A non-parametric Mann-Whitney *U* test was used to compare the RQ values of each candidate gene between

**Table 2 Primers used for qPCR**

Name of the gene	Sequence (5' - 3')	Size of PCR product (bp)	GenBank accession no.
<b>Candidate genes</b>			
LGALS3 F	CTGATTGTGCCTTATAACCTGC	100	NM_002306.3
LGALS3 R	AAGCAATTCTGTTGCATTGG		
TFF3 F	GTACGTGGGCTGTCTGC	121	NM_003226.3
TFF3 R	GATCCTGGAGTCAAAGCAGC		
DPP4 F	TGATGCTACAGCTGACAGTCG	164	NM_001935.3
DPP4 R	CTGAGCTGTTCCATATTCAGC		
CITED1 F	GCTCTGAAATGCCAACACG	174	NM_001144886.1
CITED1 R	TGGTTCCATTTGAGGCTACC		
MET F	TCTGCCTGCAATCTACAAGG	153	NM_001127500.1
MET R	AAGGTGCAGCTCTCATTTC		
CDH1 F	AGAAACAGGATGGCTGAAGG	199	NM_004360.3
CDH1 R	AGCACCTCCATGACAGACC		
CCND1 F	TGGTGAACAAGCTCAAGTGG	280	NM_053056.2
CCND1 R	ATCACTCTGGAGAGGAAGCG		
BIRC5 F	CAGCCCTTTCTCAAGGACC	152	NM_001168.2
BIRC5 R	AAGCAGAAGAACTGGGC		
<b>Housekeeping genes</b>			
PGK1 F	CTTAAGGTGCTCAACAACATGG	119	NM_000291.3
PGK1 R	ACAGGCAAGGTAATCTTCACAC		
POLR2A F	GGGTCATCTCCCAACTGGAG	164	NM_000937.4
POLR2A R	CACCAGCTTCTGCTCAATTCC		

F - forward primer, R- reverse primer

two independent groups of samples (benign vs malignant nodules; malignant nodules vs all normal tissue specimens). The statistical significance was set at  $p < 0.05$ . The receiver operating characteristic (ROC) curve was constructed and the area under the curve (AUC) was calculated to evaluate the diagnostic performance of each marker gene and biomarker model. Youden index (J) was used to define cut-off points on the ROC curves with the maximal sum of sensitivity and specificity. Multivariate logistic regression was used to examine associations between thyroid nodule histopathological status and the gene expression data as described by Laxman B et al [15]. The initial multivariate logistic regression model included all individual genes and all combinations. Stepwise backward elimination based likelihood ratio test was used to drop out insignificant terms from the initial model. The predicted probability for each sample was calculated and used as input to generate ROC curve. Leave one-out cross validation (LOOCV) as described earlier was used to validate biomarkers and combinations of them to eliminate overestimated values [15]. The statistical analyses were performed with SPSS 17.0 (SPSS, USA), Genex (Multid, Sweden) and GraphPadPrism 5 (GraphPad, USA).

## Results and discussion

### Selection of candidate marker genes

The selection of candidate marker genes was based on the functional involvement in various pathways contributing to the acquisition of malignant phenotype and/or previously reported differential expression in the malignant and benign thyroid nodules. Cyclin D1 encoded by *CCND1* is a crucial cell cycle regulator frequently upregulated at protein level in PTC [16]. *CITED1* encodes a cell cycle-dependent transcriptional cofactor involved in TGF-beta and Bmp signalling that may coordinate cellular differentiation and survival signals [17,18] and has been found to be overexpressed in PTC by expression profiling using cDNA microarrays [19], later confirmed as useful marker for FTC [6] and validated at protein level by IHC [20]. Loss of expression or function of E-cadherin - a cell-cell adhesion glycoprotein encoded by *CDH1*, has been demonstrated to contribute to the progression of various cancers, including poorly differentiated thyroid cancers by increasing proliferation, invasion and metastasis [21,22]. Overexpression of dipeptidyl-peptidase 4 encoded by *DPP4* in PTC has been demonstrated at protein level by several groups [23-25], however the mechanism how it may contribute to the development or progression of thyroid malignancies remains unknown. Survivin, an inhibitor of apoptosis encoded by *BIRC5*, has been shown to be overexpressed in a variety of cancers, including thyroid cancer, where it

contributes to uncontrolled cancer cell growth and drug resistance [26,27]. Increased expression at the mRNA and protein level of the receptor for hepatocyte growth factor encoded by *MET* has been frequently observed in PTC, follicular variant of PTC and at lesser degree in FTC where it promotes tumour progression by facilitating cell proliferation, survival, migration, invasion, and metastasis [28-30]. However, to our knowledge, so far the diagnostic value of *CDH1*, *DPP4*, *BIRC5* and *MET* mRNA levels has not been established. Galectin 3 encoded by *LGALS3* plays an important role in cell-to-cell adhesion, cell-to-matrix interactions and the regulation of apoptosis and proliferation, and its overexpression correlates with thyroid cancer aggressiveness and metastasis [31,32]. Although *TFF3* mRNA has been shown to be overexpressed in several solid cancers, such as hepatocellular carcinoma, colon and prostate cancer, and suggested to contribute the malignant behaviour and metastasis [33-36], SAGE analysis of thyroid follicular adenomas and carcinomas demonstrated that it is downregulated in FTC [37] and later this finding was confirmed in PTC and ATC by other researchers [14,38], thus validating *TFF3* downregulation as a universal marker of cancers derived from thyroid follicular cells. It encodes a small secreted protein - trefoil factor 3, which is abundantly expressed at mucosal surfaces and promotes regeneration and repair. Interestingly, recent reports demonstrate that both galectin-3 and TFF3 are involved in ciliogenesis, epithelial cell differentiation and polarity [39,40], thus suggesting a yet unexplored role of the deregulation of these processes in the development of thyroid cancer.

### mRNA expression analysis and the diagnostic performance of individual marker genes

mRNA expression analysis of the selected marker genes by RT-qPCR revealed statistically significant differences both between the normal and malignant thyroid tissues and between benign and malignant nodules for all genes except *CDH1*, thus further supporting their role the development and/or progression of thyroid cancer. *LGALS3*, *DPP4*, *MET*, *CITED1*, *CCND1* and *BIRC5* were found to be significantly upregulated while *TFF3* was downregulated in the malignant tissues (Table 3 and Figure 1). At first, all the genes were tested by ROC curve analysis as individual biomarkers, which demonstrated that *LGALS3* have the highest value for discriminating malignant from benign nodules in our sample set (AUC = 0.832,  $P < 0.0001$ ). The cut-off of *LGALS3* expression level (RQ) that discriminates benign and malignant nodules with the highest accuracy (sensitivity 90.9%, specificity 65.6%) was determined to be 0.019. The sensitivity for the detection of various subtypes of

**Table 3 Relative expression values and the diagnostic performance of the marker genes**

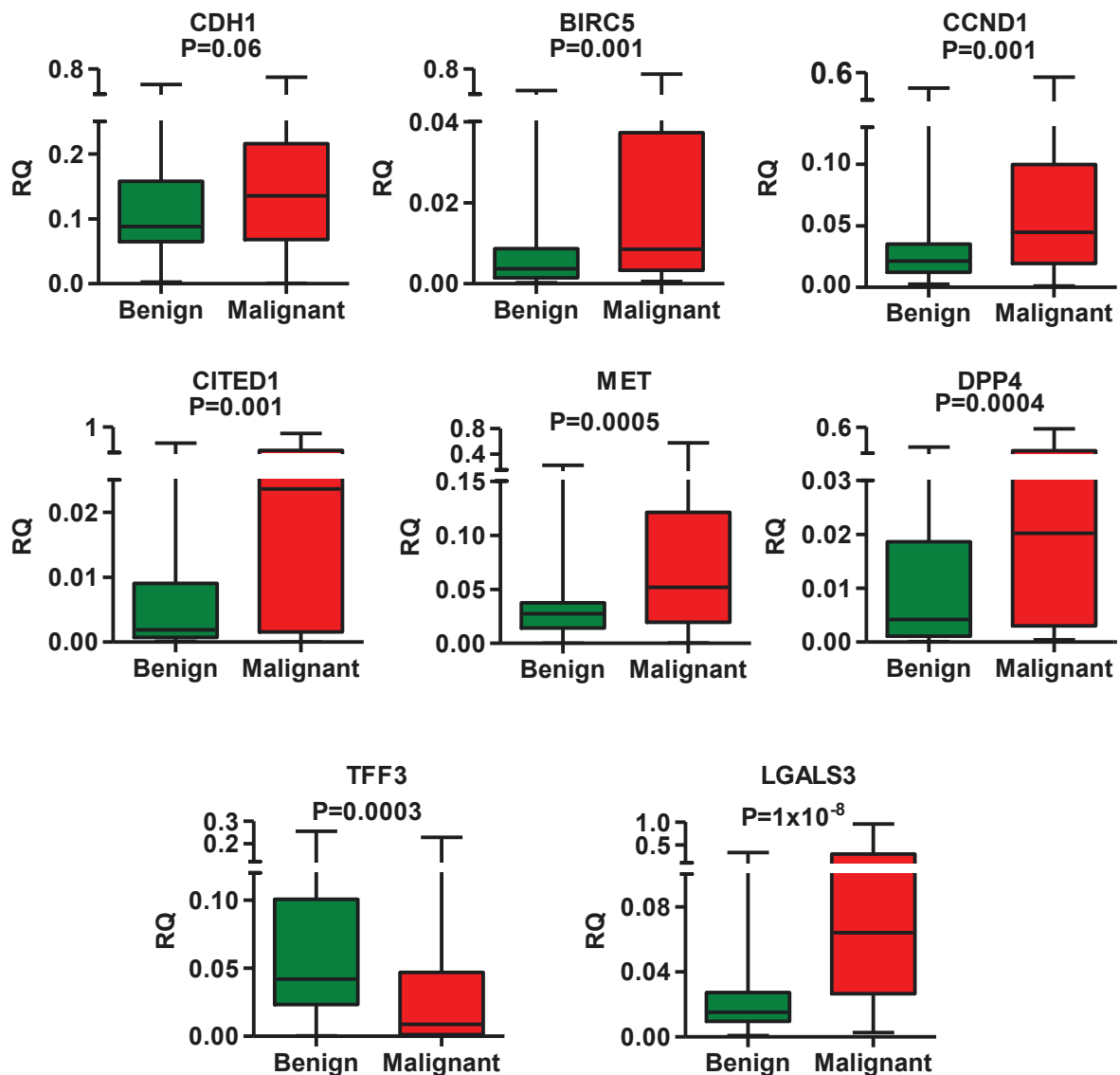
Gene, protein name	RQ values - Mean ± SEM			Mann-Whitney U test - p-values		ROC curve - AUC (Asymptotic significance)	
	Normal	Benign	Malignant	Normal vs Malignant	Benign vs Malignant	Normal vs Malignant	Benign vs Malignant
<i>LGALS3</i> , Galectin-3	0.068 ± 0.013	0.035 ± 0.007	0.248 ± 0.067	2.2 × 10 <sup>-5</sup>	1 × 10 <sup>-8</sup>	0.715 ( $< 0.0001$ )	0.832 ( $< 0.0001$ )
<i>TFF3</i> , trefoil factor 3	0.115 ± 0.011	0.066 ± 0.007	0.037 ± 0.008	4.1 × 10 <sup>-8</sup>	0.0003	0.782 ( $< 0.0001$ )	0.696 (0.001)
<i>DPP4</i> , Dipeptidyl peptidase-4	0.011 ± 0.002	0.0146 ± 0.003	0.066 ± 0.016	0.0001	0.0004	0.693 ( $< 0.0001$ )	0.695 (0.001)
<i>MET</i> , hepatocyte growth factor receptor	0.048 ± 0.005	0.033 ± 0.004	0.110 ± 0.021	0.01	0.0005	0.610 (0.037)	0.689 (0.001)
<i>CITED1</i> , Cbp/p300-interacting transactivator 1	0.010 ± 0.002	0.021 ± 0.008	0.161 ± 0.048	0.001	0.001	0.654 (0.003)	0.67 (0.003)
<i>CCND1</i> , Cyclin D1	0.037 ± 0.004	0.036 ± 0.007	0.084 ± 0.017	0.002	0.001	0.647 (0.005)	0.67 (0.003)
<i>BIRC5</i> , Survivin	0.025 ± 0.011	0.009 ± 0.002	0.045 ± 0.018	0.0001	0.001	0.694 ( $< 0.0001$ )	0.668 (0.003)
<i>CDH1</i> , E-cadherin	0.171 ± 0.013	0.117 ± 0.009	0.158 ± 0.018	0.2	0.06	0.464 (0.497)	0.525 (0.697)

thyroid cancers was mutually comparable (Figure 2C) - 29 out of 33 PTC, 4 of 5 MTC and all FTC and ATC specimens were classified correctly thus confirming that *LGALS3* may serve as a universal diagnostic marker of the thyroid malignancies. However, 22 of 61 benign nodules were misclassified using this cut-off. As the ROC curve was constructed using all the samples, it may lead to overestimation of the AUC, therefore we next used LOOCV to validate it. The AUC for LOOCV *LGALS3* model dropped to 0.783, nevertheless it still outperformed the other marker genes. Although previously the diagnostic utility of *LGALS3* mRNA expression level has been questioned, our data show that the diagnostic accuracy of mRNA quantification is comparable to that reported in immunohistochemical studies [32]. The next marker genes with the highest diagnostic value in our sample set were *TFF3* (AUC = 0.696, P = 0.001) and *DPP4* (AUC = 0.695 and P = 0.001). The optimal RQ cut-off for *TFF3* was determined to be 0.011 allowing the discrimination between benign and malignant nodules with sensitivity 54.6% and specificity 90.2%, while for *DPP4* the cut-off was 0.027 with sensitivity 45.6% and specificity 88.5%.

Interestingly, in 3 PTC and 2 MTC cases none or only one of the genes analysed were differentially expressed between cancerous and relatively normal tissues in specimens and they were consistently misclassified as benign nodules. These specimens are likely to represent a distinct molecular subtype of thyroid cancer, in which different molecular pathways predominate and therefore next these samples will be subjected to gene expression profiling using cDNA microarrays.

#### Development of multiplex biomarker model

To determine if a combination of markers could outperform the single biomarkers we systematically searched for two-marker ratios or two-marker sums, and determined their diagnostic performance by ROC curve analysis. In total, 27 two-marker combinations were evaluated, however only one of them - *LGALS3* and *BIRC5* RQ sum showed higher AUC for discriminating benign and malignant nodules than the best individual marker gene (AUC = 0.841, P < 0.0001 and AUC = 0.809, P < 0.0001 for the LOOCV model) (Figure 2A,B). At the best cut-off, the sensitivity was 79.5% and specificity was 78.7%. As shown in Figure 2D, this two-gene combination could detect all FTC and ATC cases but failed to detect 7 of 33 PTCs (21%) and 2 of 5 MTCs (40%) thus suggesting it may have lower value for diagnosing MTC, however this finding should be verified in larger cohort of patients. So far, the best reported two-marker set is *TFF3/LGALS3* ratio that has been shown to discriminate follicular carcinomas from follicular adenomas with 72.4% sensitivity and 83.3% specificity (or 80% and 91.5%, when the pathologically questionable cases were excluded) [41]. In our sample set it showed similar performance (AUC = 0.758, P < 0.0001, sensitivity 72.7%, specificity 85.3%). However, although it could improve the specificity, the overall accuracy of *TFF3/LGALS3* ratio was lower than that of *LGALS3* alone. Moreover, as *TFF3* is downregulated in cancer cells, the development of assay that is based on the measurement of its expression level in FNA biopsies consisting of a mixture of different cell types is technically more challenging than the assay

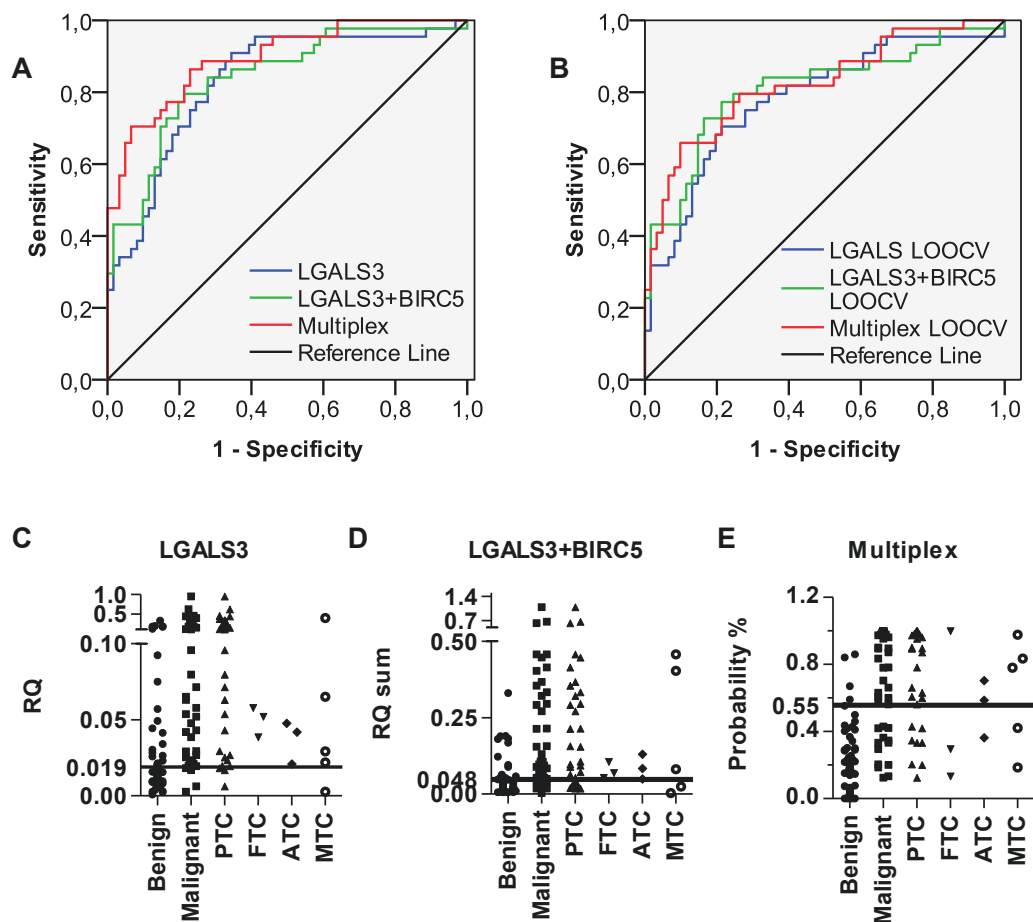


**Figure 1** Median mRNA expression levels of the candidate marker genes in groups of 61 benign and 44 malignant thyroid nodules represented by box plots. The relative quantity (RQ) of each marker gene in each sample was calculated using  $2^{-\Delta CT}$  method and normalised by using normalisation factor obtained by analysis of the two most stable reference genes ( $RQ = (2^{-\Delta CT})/NF$ ). Boxes represent the 25<sup>th</sup> and 75<sup>th</sup> percentiles; whiskers represent the minimum and maximum values. Statistical significance was calculated using Mann-Whitney *U* test.

based on overexpressed genes, therefore *LGALS3* and *BIRC5* RQ sum seems to be more suitable for the clinical application.

Next, we applied a multivariate logistic regression analysis to define a multiple marker set that could improve the diagnostic performance over single markers or two-marker combinations. Similar approach has been previously successfully used by Laxman B et al to develop a multiplex biomarker model for the detection of prostate cancer [15]. This resulted in the multiplex model that included *LGALS3* and *BIRC5* RQ sum, *TFF3* and *LGALS3* RQ ratio, *TFF3* and *CCND1* RQ ratio, ratio

between *TFF3* RQ and *MET*, *CITED1* RQ sum, ratio between *TFF3* RQ and *MET*, *BIRC5* RQ sum and ratio between *TFF3* RQ and *CCND1*, *BIRC5* RQ sum. Although the overall performance of the model (AUC = 0.895,  $P < 0.0001$  and AUC = 0.823,  $P < 0.0001$  for the LOOCV model) was only slightly improved over *LGALS3* alone or *LGALS3* and *BIRC5* RQ sum, at the best cut-off, this model shows 70.5% sensitivity and 93.4% specificity and as shown in Figure 2A it has considerably higher specificity that is a clear requirement for the development of clinically applicable biomarker assay.



Test Result Variable(s)						Asymptotic 95% Confidence Interval	
	Area	Std. Error	Asymptotic Sig.	Sens.	Spec.	Lower Bound	Upper Bound
LGALS3	0.832	0.041	<0.0001	90.9%	65.6%	0.752	0.913
LGALS3 LOOCV	0.783	0.047	<0.0001	70.5%	78.7%	0.691	0.875
LGALS3+BIRC5	0.841	0.04	<0.0001	79.5%	78.7%	0.763	0.919
LGALS3+BIRC5 LOOCV	0.809	0.046	<0.0001	77.3%	78.7%	0.720	0.898
Multiplex	0.895	0.031	<0.0001	70.5%	93.4%	0.834	0.955
Multiplex LOOCV	0.823	0.042	<0.0001	65.9%	90.2%	0.74	0.906

**Figure 2** The diagnostic performance of the best individual marker, two-marker set and multiplex biomarker model. A and B represents ROC curves of the best individual marker, two-marker set and multiplex model and their LOOCV ROC curves, respectively. C, D, E - Scatter dot plot of LGALS3 RQ, LGALS3 and BIRC5 RQ sum and Multiplex model with the best cut off shows how many benign nodules, malignant nodules and thyroid cancer subtypes are misclassified. In table are the best individual marker, two-marker set and multiplex biomarker model and their LOOCV model ROC curve AUC data and sensitivity and specificity at the best cut-off point.

The future efforts will be focused on adding marker genes to the multiplex model that would enable detecting other molecular subtypes of thyroid cancer, testing the feasibility of the biomarker assay in FNA biopsies and validating the assay in a large multicenter clinical trial.

## Conclusions

mRNA expression analysis of 8 candidate marker genes - *BIRC5*, *CCND1*, *CDH1*, *CITED1*, *DPP4*, *LGALS3*, *MET* and *TFF3* in paired nodular and relatively normal thyroid tissue specimens of 105 consecutive patients undergoing thyroid surgery demonstrated that all of them except *CDH1* are differentially expressed between the normal and malignant thyroid tissues and between benign and malignant nodules, and *LGALS3* had the highest diagnostic value for the discrimination of malignant and benign thyroid nodules on a single gene basis. An application of multivariate logistic regression analysis resulted in the generation of a multiplex biomarker model based on *LGALS3*, *BIRC5*, *TFF3*, *CCND1*, *MET* and *CITED1* that outperformed a single marker or two marker gene-based models by increasing the specificity, which is a prerequisite for the development of clinically applicable biomarker assay. The next step will be to test the feasibility of this assay on FNA biopsies and to validate it in a larger cohort of patients.

## Acknowledgements and funding

This study was supported in parts by funds from Latvian Council of Science, grant number 09.1310, European Social Fund project No. 2009/0204/1DP/1.1.1.2.0/09/APIA/VIAA/150 and individual fellowships from European Social Fund No. 10/172/31. The publishing costs are covered by ERDF project No. 2DP/2.1.1.2.0/10/APIA/VIAA/004.

## Author details

<sup>1</sup>Faculty of Medicine, University of Latvia, Raina Bulvaris 19, Riga, LV1586, Latvia. <sup>2</sup>Department of Molecular Biology, Latvian Biomedical Research and Study Centre, Ratsupites 1, Riga, LV1067, Latvia. <sup>3</sup>Surgery Clinics, Pauls Stradins Clinical University Hospital, Pilsonu 13, Riga, LV1002, Latvia. <sup>4</sup>Oncology department, Riga Eastern Clinical University Hospital, Hipokrata 2, Riga, LV1038, Latvia.

## Authors' contributions

KD participated in the designing of the study, collecting of clinical data and drafting the manuscript. AA carried out the gene expression analysis, performed the statistical analysis and helped to draft the manuscript. JV, ZN and JT performed thyroid surgery and collected the clinical material. DA participated in the gene expression analysis. AL participated in the design of the study and drafted the manuscript. VP designed and coordinated the study and helped to draft the manuscript. All authors read and approved the final manuscript.

## Competing interests

The authors declare that they have no competing interests.

Received: 12 April 2011 Accepted: 27 June 2011  
Published: 27 June 2011

## References

- Gharib H, Papini E, Valcavi R, Baskin HJ, Crescenzi A, Dottorini ME, et al: American Association of Clinical Endocrinologists and Associazione Medici Endocrinologi medical guidelines for clinical practice for the diagnosis and management of thyroid nodules. *Endocr Pract* 2006, **12**:63-102.
- Liu S, Semenciw R, Ugnat AM, Mao Y: Increasing thyroid cancer incidence in Canada, 1970-1996: time trends and age-period-cohort effects. *Br J Cancer* 2001, **85**:1335-1339.
- Ferlay J, Parkin DM, Steliarova-Foucher E: Estimates of cancer incidence and mortality in Europe in 2008. *Eur J Cancer* 2010, **46**:765-781.
- Kondo T, Ezzat S, Asa SL: Pathogenetic mechanisms in thyroid follicular-cell neoplasia. *Nat Rev Cancer* 2006, **6**:292-306.
- Lloyd RV, Erickson LA, Casey MB, Lam KY, Lohse CM, Asa SL, et al: Observer variation in the diagnosis of follicular variant of papillary thyroid carcinoma. *Am J Surg Pathol* 2004, **28**:1336-1340.
- Foukakis T, Gusnanto A, Au AY, Hoog A, Lui WO, Larsson C, et al: A PCR-based expression signature of malignancy in follicular thyroid tumors. *Endocr Relat Cancer* 2007, **14**:381-391.
- Barden CB, Shister KW, Zhu B, Guiter G, Greenblatt DY, Zeiger MA, et al: Classification of follicular thyroid tumors by molecular signature: results of gene profiling. *Clin Cancer Res* 2003, **9**:1792-1800.
- Finley DJ, Arora N, Zhu B, Gallagher L, Fahey TJ III: Molecular profiling distinguishes papillary carcinoma from benign thyroid nodules. *J Clin Endocrinol Metab* 2004, **89**:3214-3223.
- Finley DJ, Lubitz CC, Wei C, Zhu B, Fahey TJ III: Advancing the molecular diagnosis of thyroid nodules: defining benign lesions by molecular profiling. *Thyroid* 2005, **15**:562-568.
- Nardone HC, Ziober AF, LiVolsi VA, Mandel SJ, Baloch ZW, Weber RS, et al: c-Met expression in tall cell variant papillary carcinoma of the thyroid. *Cancer* 2003, **98**:1386-1393.
- Finley DJ, Zhu B, Barden CB, Fahey TJ III: Discrimination of benign and malignant thyroid nodules by molecular profiling. *Ann Surg* 2004, **240**:425-436.
- Jarzab B, Wiench M, Fajarewicz K, Simek K, Jarzab M, Oczko-Wojciechowska M, et al: Gene expression profile of papillary thyroid cancer: sources of variability and diagnostic implications. *Cancer Res* 2005, **65**:1587-1597.
- Cerutti JM, Delcelo R, Amadei MJ, Nakabashi C, Maciel RM, Peterson B, et al: A preoperative diagnostic test that distinguishes benign from malignant thyroid carcinoma based on gene expression. *J Clin Invest* 2004, **113**:1234-1242.
- Krause K, Eszlinger M, Gimm O, Karger S, Engelhardt C, Dralle H, et al: TFF3-based candidate gene discrimination of benign and malignant thyroid tumors in a region with borderline iodine deficiency. *J Clin Endocrinol Metab* 2008, **93**:1390-1393.
- Laxman B, Morris DS, Yu J, Siddiqui J, Cao J, Mehra R, et al: A first-generation multiplex biomarker analysis of urine for the early detection of prostate cancer. *Cancer Res* 2008, **68**:645-649.
- Lee SH, Lee JK, Jin SM, Lee KC, Sohn JH, Chae SW, et al: Expression of cell-cycle regulators (cyclin D1, cyclin E, p27kip1, p57kip2) in papillary thyroid carcinoma. *Otolaryngol Head Neck Surg* 2010, **142**:332-337.
- Plisov S, Tsang M, Shi G, Boyle S, Yoshino K, Dunwoodie SL, et al: Cited1 is a bifunctional transcriptional cofactor that regulates early nephronic patterning. *J Am Soc Nephrol* 2005, **16**:1632-1644.
- Shi G, Boyle SC, Sparrow DB, Dunwoodie SL, Shioda T, de Caestecker MP: The transcriptional activity of CITED1 is regulated by phosphorylation in a cell cycle-dependent manner. *J Biol Chem* 2006, **281**:27426-27435.
- Huang Y, Prasad M, Lemon WJ, Hampel H, Wright FA, Kornacker K, et al: Gene expression in papillary thyroid carcinoma reveals highly consistent profiles. *Proc Natl Acad Sci USA* 2001, **98**:15044-15049.
- Liu YY, Morreau H, Kievit J, Romijn JA, Carrasco N, Smit JW: Combined immunostaining with galectin-3, fibronectin-1, CITED-1, Hector Battifora mesothelial-1, cytokeratin-19, peroxisome proliferator-activated receptor- $\gamma$ , and sodium/iodide symporter antibodies for the differential diagnosis of non-medullary thyroid carcinoma. *Eur J Endocrinol* 2008, **158**:375-384.
- Mitselou A, Ioachim E, Peschos D, Charalabopoulos K, Michael M, Agnantis NJ, et al: E-cadherin adhesion molecule and syndecan-1 expression in various thyroid pathologies. *Exp Oncol* 2007, **29**:54-60.



22. Rocha AS, Soares P, Fonseca E, Cameselle-Teijeiro J, Oliveira MC, Sobrinho-Simoes M: **E-cadherin loss rather than beta-catenin alterations is a common feature of poorly differentiated thyroid carcinomas.** *Histopathology* 2003, **42**:580-587.
23. de MC, Savchenko V, Giorgi R, Sebag F, Henry JF: **Utility of malignancy markers in fine-needle aspiration cytology of thyroid nodules: comparison of Hector Battifora mesothelial antigen-1, thyroid peroxidase and dipeptidyl aminopeptidase IV.** *Br J Cancer* 2008, **98**:818-823.
24. Aratake Y, Umeki K, Kiyoyama K, Hinoura Y, Sato S, Ohno A, *et al*: **Diagnostic utility of galectin-3 and CD26/DPPIV as preoperative diagnostic markers for thyroid nodules.** *Diagn Cytopathol* 2002, **26**:366-372.
25. Kholova I, Ludvikova M, Ryska A, Hanzelkova Z, Cap J, Pecan L, *et al*: **Immunohistochemical detection of dipeptidyl peptidase IV (CD 26) in thyroid neoplasia using biotinylated tyramine amplification.** *Neoplasma* 2003, **50**:159-164.
26. Cheung CH, Cheng L, Chang KY, Chen HH, Chang JY: **Investigations of survivin: the past, present and future.** *Front Biosci* 2011, **16**:952-961.
27. Zhang HY, Meng X, Du ZX, Fang CQ, Liu GL, Wang HQ, *et al*: **Significance of survivin, caspase-3, and VEGF expression in thyroid carcinoma.** *Clin Exp Med* 2009, **9**:207-213.
28. Steffan JJ, Coleman DT, Cardelli JA: **The HGF-Met Signaling Axis: Emerging Themes and Targets of Inhibition.** *Curr Protein Pept Sci* 2011, **12**:12-22.
29. Ippolito A, Vella V, La Rosa GL, Pellegriti G, Vigneri R, Belfiore A: **Immunostaining for Met/HGF receptor may be useful to identify malignancies in thyroid lesions classified suspicious at fine-needle aspiration biopsy.** *Thyroid* 2001, **11**:783-787.
30. Wasenius VM, Hemmer S, Kettunen E, Knuutila S, Franssila K, Joensuu H: **Hepatocyte growth factor receptor, matrix metalloproteinase-11, tissue inhibitor of metalloproteinase-1, and fibronectin are up-regulated in papillary thyroid carcinoma: a cDNA and tissue microarray study.** *Clin Cancer Res* 2003, **9**:68-75.
31. Htwe TT, Karim N, Wong J, Jahanfar S, Mansur MA: **Differential expression of galectin-3 in advancing thyroid cancer cells: a clue toward understanding tumour progression and metastasis.** *Singapore Med J* 2010, **51**:856-859.
32. Chiu CG, Strugnell SS, Griffith OL, Jones SJ, Gown AM, Walker B, *et al*: **Diagnostic utility of galectin-3 in thyroid cancer.** *Am J Pathol* 2010, **176**:2067-2081.
33. Babyatsky M, Lin J, Yio X, Chen A, Zhang JY, Zheng Y, *et al*: **Trefoil factor-3 expression in human colon cancer liver metastasis.** *Clin Exp Metastasis* 2009, **26**:143-151.
34. Okada H, Kimura MT, Tan D, Fujiwara K, Igarashi J, Makuuchi M, *et al*: **Frequent trefoil factor 3 (TFF3) overexpression and promoter hypomethylation in mouse and human hepatocellular carcinomas.** *Int J Oncol* 2005, **26**:369-377.
35. Yio X, Zhang JY, Babyatsky M, Chen A, Lin J, Fan QX, *et al*: **Trefoil factor family-3 is associated with aggressive behavior of colon cancer cells.** *Clin Exp Metastasis* 2005, **22**:157-165.
36. Vestergaard EM, Nexø E, Tørring N, Borre M, Orntoft TF, Sørensen KD: **Promoter hypomethylation and upregulation of trefoil factors in prostate cancer.** *Int J Cancer* 2010, **127**:1857-1865.
37. Takano T, Miyauchi A, Yoshida H, Kuma K, Amino N: **High-throughput differential screening of mRNAs by serial analysis of gene expression: decreased expression of trefoil factor 3 mRNA in thyroid follicular carcinomas.** *Br J Cancer* 2004, **90**:1600-1605.
38. Taniguchi K, Takano T, Miyauchi A, Koizumi K, Ito Y, Takamura Y, *et al*: **Differentiation of follicular thyroid adenoma from carcinoma by means of gene expression profiling with adapter-tagged competitive polymerase chain reaction.** *Oncology* 2005, **69**:428-435.
39. Koch A, Poirier F, Jacob R, Delacour D: **Galectin-3, a novel centrosome-associated protein, required for epithelial morphogenesis.** *Mol Biol Cell* 2010, **21**:219-231.
40. LeSimple P, van S, Buisine MP, Copin MC, Hinz M, Hoffmann W, *et al*: **Trefoil factor family 3 peptide promotes human airway epithelial ciliated cell differentiation.** *Am J Respir Cell Mol Biol* 2007, **36**:296-303.
41. Takano T, Miyauchi A, Yoshida H, Kuma K, Amino N: **Decreased relative expression level of trefoil factor 3 mRNA to galectin-3 mRNA distinguishes thyroid follicular carcinoma from adenoma.** *Cancer Lett* 2005, **219**:91-96.

doi:10.1186/1756-6614-4-11

**Cite this article as:** Ducena *et al*: Validity of multiplex biomarker model of 6 genes for the differential diagnosis of thyroid nodules. *Thyroid Research* 2011 **4**:11.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

