Oncogene (2007), 1–10
© 2007 Nature Publishing Group All rights reserved 0950-9232/07 $30.00
www.nature.com/onc

ORIGINAL ARTICLE

Gene expression and the biological phenotype of papillary thyroid carcinomas

L Delys[1,5], V Detours[1,5], B Franc[2], G Thomas[3], T Bogdanova[4], M Tronko[4], F Libert[1], JE Dumont[1], and C Maenhaut[1]

[1]*Institute of Interdisciplinary Research, School of Medicine, Université Libre de Bruxelles, Campus Erasme, Brussels, Belgium;* [2]*Service d'Anatomie et de Cytologie Pathologiques, Faculté de Médecine Paris Ile de France Ouest, Hôpital Ambroise Paré (APHP), Université Versailles Saint-Quentin en Yvelines, Boulogne, France;* [3]*South West Wales Cancer Institute/Swansea Clinical School, Singleton Hospital, Swansea, UK and* [4]*Institute of Endocrinology and Metabolism, Kiev, Ukraine*

The purpose of this paper is to correlate the molecular phenotype of papillary thyroid carcinoma (PTC) to their biological pathology. We hybridized 26 PTC on microarrays and showed that nearly 44% of the transcriptome was regulated in these tumors. We then combined our data set with two published PTC microarray studies to produce a platform- and study-independent list of PTC-associated genes. We further confirmed the mRNA regulation of 15 genes from this list by quantitative reverse transcription–PCR. Analysis of this list with statistical tools led to several conclusions: (1) there is a change in cell population with an increased expression of genes involved in the immune response, reflecting lymphocyte infiltration in the tumor compared to the normal tissue. (2) The c-jun N-terminal kinase pathway is activated by overexpression of its components. (3) The activation of ERKK1/2 by genetic alterations is supplemented by activation of the epidermal growth factor but not of the insulin-like growth factor signaling pathway. (4) There is a downregulation of immediate early genes. (5) We observed an overexpression of many proteases in accordance with tumor remodeling, and suggested a probable role of S100 proteins and annexin A2 in this process. (6) Numerous overexpressed genes favor the hypothesis of a collective migration mode of tumor cells.
Oncogene advance online publcation, 9 July 2007; doi:10.1038/sj.onc.1210588

Keywords: cancer; thyroid; microarrays; molecular phenotype

Introduction

Papillary thyroid carcinoma (PTC) is the most frequent endocrine malignancy in human and represents up to 80% of all malignant thyroid tumors. PTC is usually biologically indolent and has an overall 5- to 10-year survival rate of 80–95%. Lymph node metastasis is commonly found in patients with PTC contrasting with a low rate of distant metastases (Gimm, 2001). Their diagnosis is based on the presence of a number of different features, not all of which need to be present in the same lesion, such as papillary architecture, the presence of psammoma bodies and characteristic nuclear features. PTC also often displays lymphocytic infiltration and fibrosis. Different pathological subtypes of PTC have been described, including the classical, follicular, solid and tall-cell variants (Gimm, 2001; Livolsi *et al*., 2004).

After the Chernobyl power plant explosion, an unusual number of childhood thyroid cancers were observed in Belarus and Ukraine, with an incidence 10- to 100-fold higher than in the rest of Europe. These cancers have been described mostly as PTC and we have shown previously that they belong to the same entity as sporadic PTC (Detours *et al*., 2005).

Several lines of evidence point to the causative role of chromosomal rearrangements and point mutations in the pathogenesis of PTC leading to constitutively activated effectors along the RAS/RAF/MEK/ERK signaling pathway. Two membrane tyrosine kinase receptors, RET and less frequently NTRK1, are commonly found rearranged in PTC (Alberti *et al*., 2003). AKAP9-BRAF rearrangement has also been reported in some post-Chernobyl PTC of short latency (Ciampi and Nikiforov, 2005). These rearranged proteins lead to the constitutive kinase activity of the oncogene. Besides chromosomal rearrangements, BRAF point mutations have been described in PTC and in other human cancers (Ciampi and Nikiforov, 2005). These point mutations produce a protein with constitutive serine–threonine kinase activity. Activating mutations of RAS have been also found in the follicular variant of PTC leading to constitutive activation of the ERK signaling pathway.

Current microarray studies on tumors are mostly used to define diagnostic and prognostic signatures. PTC gene expression analyses have for example provided tumor type signatures and views on tumor cell

Correspondence: Dr L Delys and Dr C Maenhaut, Institute of Interdisciplinary Research, School of Medicine, Free University of Brussels, Campus Erasme, Route de Lennik 808, Brussels B-1070, Belgium.
E-mails: laurent.delys@ulb.ac.be and cmaenhau@ulb.ac.be
[5]These authors contributed equally to this work.
Received 29 November 2006; revised 3 April 2007; accepted 11 May 2007

metabolism (Baris *et al*., 2005; Giordano *et al*., 2005). However, most of these studies rarely analyse in details the biological function of the regulated genes and their potential implication in tumor initiation and progression. On the other hand, most of the proteomic studies are focalized on only one or two particular proteins, without taking into account the other proteins involved in the same cascade or the same process. In this paper, we performed microarray experiments on sporadic and post-Chernobyl PTC and combined our data with two other independent microarray data sets (Huang *et al*., 2001; Jarzab *et al*., 2005) to obtain for the first time for PTC a cross-validated regulated gene list. As the transcription level of an mRNA usually reflects the level of the corresponding protein, we used this gene list to attempt to give a general view of the regulation of different signaling pathways and processes and correlated these results to the physiopathology of the PTC.

Results

The expression of thousands of genes is altered in PTC with a high level of statistical evidence

RNA samples were collected from PTC tumors and non-tumor tissue counterparts from 26 patients. Tumor and patient-matched non-tumor RNA samples were cohybridized on Agilent Human 1a cDNA microarrays (see Materials and methods). The 12 000 clones spotted on the Agilent slides were then interrogated for consistent up- or downregulation across patients. The level of statistical evidence for regulation was estimated with the Significance Analysis of Microarray (SAM) procedure. This procedure avoids normality assumptions, and handles efficiently the fact that thousands statistical tests are conducted at once by estimating statistical significance in term of q-values. SAM reports that 44.5% of the genes shows statistical evidence ($q<0.05$) for differential expression. Running SAM on the expression profiles of 16 PTC from Jarzab *et al*. (2005) gives 40.3% with $q<0.05$. Most of the significantly regulated genes are regulated at levels well below two fold in both studies. Only 6.7% of the genes are below the 0.05 confidence threshold in the data set of Huang *et al*. (2001). A reanalysis of this later study (Pavlidis *et al*., 2003) suggests that its small size (eight patients), hence low statistical power, could in part explain the discrepancy with our and Jarzab *et al*. (2005) studies. Thus, we concluded that a very large fraction of the transcriptome is significantly regulated in PTC with respect to healthy tissue counterparts, which would be in keeping with the greatly altered morphology of the tumor relative to normal thyroid.

A reliable list of genes regulated in PTC

To produce a reliable list of regulated genes in PTC, we combined our data set with two independent PTC microarray data sets, from Jarzab *et al*. (2005) and Huang *et al*. (2001). They contain 16 and 8 pairs of PTC tumors and patient-matched normal tissues, respectively. Thus, our global analysis was based on 50 tumors derived from patients of different ages presenting various histological variants and genetic alterations. Moreover, the three data sets were produced independently on different microarray platforms (cDNA and oligonucleotide chips) with different protocols. We compiled all these data to create a reference list of genes modulated in PTC. We chose to include a gene in our gene list only when it was modulated in two of the data sets at least, with a minimum ratio of 1.5 and a maximum q-value of 0.05. This q-value enabled us to select only genes regulated in most of the PTC, regardless of the histological variant or the age of the patients. The resulting list is composed of 451 up- and 233 downregulated genes representing the general molecular phenotype of PTC (Supplementary Table 1). The probability of obtaining this large extent of overlap between the three data sets was computed using a resampling approach (see Supplementary Information): it is very low, $P<10^{-5}$. This high significance follows obviously from the excellent agreement between the data sets: correlation between our data and Jarzab's is 0.77, 0.65 with Huang and 0.70 between Huang and Jarzab data sets. Note that this gene list comes from three publicly available data sets and is therefore open to independent recalculation.

Confirmation of the modulation of selected genes by real time reverse transcription–PCR using two non-modulated genes confirmed in PTC

To select adequate normalization genes for PTC, we first investigated different candidates and identified *NEDD8* (neuronal precursor cell expressed, developmentally downregulated 8) and *TTC1* (tetratricopeptide repeat domain 1) as a stable combination of non-regulated genes to normalize real-time reverse transcription (RT)–PCR measurements in PTC (see Materials and methods).

Using *NEDD8* and *TTC1* for normalization, 11 upregulated genes (*ANXA1*, *CDH3*, *CLDN1*, *DUSP5*, *GPX1*, *HMGA2*, *NELL2*, *NRCAM*, *SLIT1*, *THBS2*, *TNC*) and four downregulated genes (*BCL2*, *EGR1*, *EGR2*, *FLRT2*) identified by microarray were confirmed by real-time RT–PCR (Figure 2). This also included two genes for which the ratio tumor/normal was smaller than two for upregulated genes (*GPX1*, *THBS2*) in our microarray data. In addition, two downregulated genes from the Jarzab's data set (*MAFB* and *DGKI*) but not present in our or in the Huang's data set were also confirmed (not shown). A Pearson correlation of 89% was obtained between microarrays and real-time RT–PCR results.

In addition, representativity of our list is supported by several agreements:

(1) Hybridization of the same pair of samples on two Agilent microarray slides gave a correlation of 95% (not shown).
(2) Previously reported data on RNA regulation of many genes have been confirmed. For example, of the 26 genes regulated in PTC and tabulated from the literature in a review article, 20 are in our list (Kondo *et al*., 2006). Confirmation at protein level

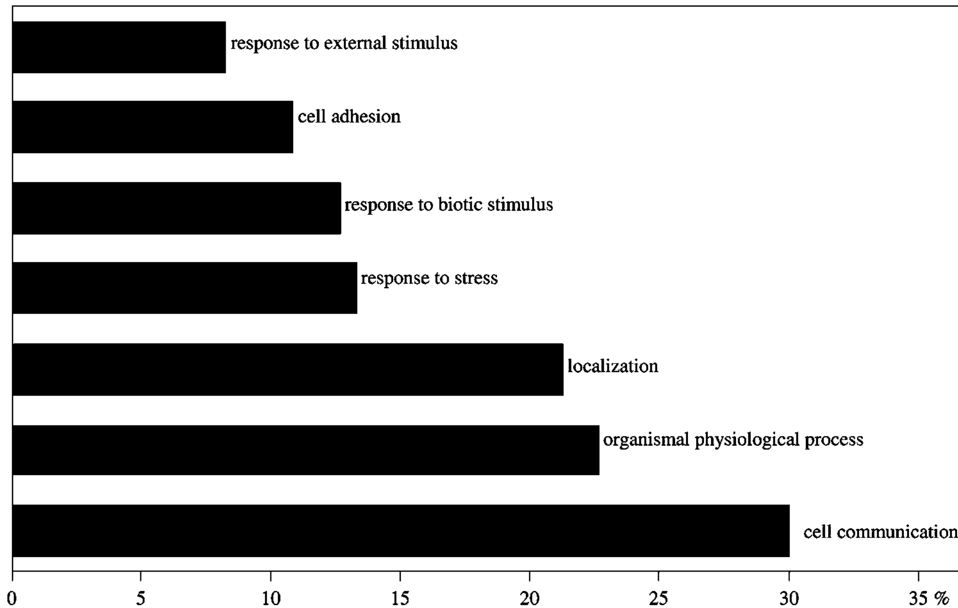

Figure 1 Gene function distribution (in %) altered in PTC according to the Gene Ontology. PTC, papillary thyroid carcinoma.

Table 1 GO categories and statistical significance following the analysis of the PTC-regulated gene list with DAVID software

GO identifier	*GO name*	P-*value*
GO:0006955	Immune response	0.0000066
GO:0005125	Cytokine activity	0.000012
GO:0008009	Chemokine activity	0.0002
GO:0006590	Thyroid hormone generation	0.014
GO:0007173	EGFR signaling pathway	0.046
GO:0007257	Activation of JNK activity	0.019
GO:0017017	MAP kinase phosphatase activity	0.042
GO:0031012	Extracellular matrix	4.9E-11
GO:0008233	Peptidase activity	0.0049
GO:0030414	Protease inhibitor activity	0.021
GO:0043256	Laminin complex	0.047
GO:0005581	Collagen	0.00076
GO:0016337	Cell–cell adhesion	0.0034
GO:0008305	Integrin complex	0.063

Abbreviations: DAVID, database for annotation, visualization and integrated discovery; EGFR, epidermal growth factor receptor; GO, Gene Ontology; JNK, c-jun N-terminal kinase; MAP kinase, mitogen-activated protein kinase; PTC, papillary thyroid carcinoma.

of numerous genes present in our gene list are also found in the literature (for example SPP1, TGFA, ERBB3, HRG, PLAU, MMP1, TIMP1, S100A4, ICAM1, …). Moreover, the modulation of three proteins (CDH2, CDH3 and ANXA1) was confirmed by our group (data not shown).

Classification of our gene list in Gene Ontology categories using the DAVID software

To assess the most representative biological activities present in our gene list, we used the statistical methods from the DAVID (database for annotation, visualization and integrated discovery) software (Dennis *et al.*, 2003; Hosack *et al.*, 2003), which finds the most represented functions according to the Gene Ontology (GO) annotations. As shown in Figure 1, the main global biological processes (GO level 2) altered in PTCs were cell communication, organismal physiological process, localization, cell adhesion and diverse responses to stimulus and stress.

We then analysed in details the GO categories detected by DAVID with a *P*-value <0.05 to relate some aspects of the PTC phenotype with its gene expression profile (Table 1).

Discussion

In this study, we performed cDNA microarrays on 26 PTC and showed that more than 40% of the transcriptome was regulated in PTC. Several factors may contribute to this, including alteration of thyrocyte metabolism and intracellular signaling, changes in relative cell populations, for example, lymphocyte infiltration (Jarzab *et al.*, 2005) and adaptation of adjacent tissues to the neighboring tumors. Then, we combined our data set with two other data sets to generate a cross-validated list of regulated genes in PTC. We used this gene list to describe the common molecular phenotype of PTC and to relate it to the biology of the tumor.

Change in relative cell populations has a major impact on the difference in gene expression between normal and tumor tissues

Our global analyses revealed many genes involved in the response to diverse stimulus and to stress in PTC. This resulted in many highly significant GO categories related to the immune response, including the immune response category (Table 1). Most of the genes present in these categories were overexpressed (see Supplementary Table 2). This suggested heavy infiltrations of the

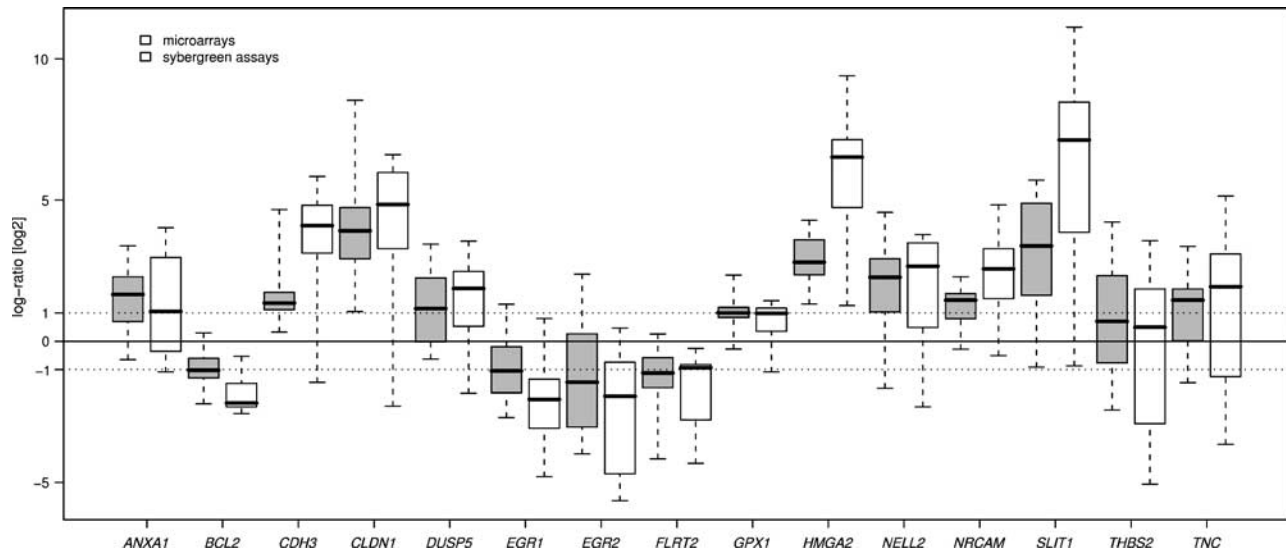

Figure 2 Comparison of differential gene expression data obtained by microarrays and real-time RT–PCR. The upper and lower limits of each box stand for the upper and the lower quartiles, respectively; bold lines represent medians; whiskers represent extreme measurements. Regulation of *CDH3*, *CLDN1*, *HMGA2*, *NRCAM*, *SLIT1* were confirmed on 20 tumor/non-tumor pairs of PTC whereas 11 pairs of PTC were used to confirm *ANXA1*, *BCL2*, *DUSP5*, *EGR1*, *EGR2*, *FLRT2*, *GPX1*, *NELL2*, *THBS2* and *TNC*. PTC, papillary thyroid carcinoma; RT–PCR, reverse transcription–PCR.

tumors by immune cells. To support this hypothesis, we ran the Gene Set Enrichment Analysis (Subramanian *et al*., 2005) to verify that genes present in the 'immune response' GO category were significantly overrepresented in lymphocytes-infiltrated versus non or low infiltrated tumors. The result was statistically significant with $P<1/2000$ (data not shown). This adds weight to the conclusions of Jarzab *et al*. (2005).

PTC is often associated with striking chronic inflammatory reaction (Livolsi *et al*., 2004). It is consistent with the experimental observation that expression of the RET/PTC-rearranged human gene in mice thyroid leads to a tumor with a strong immune response and inflammation (Powell *et al*., 2003). This suggests that the tumor cells themselves induce the inflammation rather than the reverse, like in hepatitis generated hepatocarcinomas. The induction in the tumor of many inflammatory cytokines could account for the inflammation. Indeed, RET/PTC induces such cytokines in PCCl3 cells and in human thyrocytes (Borrello *et al*., 2005). Our gene list revealed that different cytokines and chemokines were significantly altered in PTC (Table 1) with a trend in upregulation (Supplementary Table 3). The role of this inflammation on tumor progression remains to be determined. One would expect an antitumoral role, but inflammation may favor tumor progression (de Visser *et al*., 2006) and macrophages may be partners for tumor cell migration, invasion and metastasis. This raises the question of whether anti-inflammatory treatment would be beneficial for these tumors.

Other changes may reflect gene regulation in the cancer cells themselves. However, the possibility that some upregulation of gene expression may at least in part reflect the increased proportion of inflammatory or even endothelial cells (data not shown) should be recognized. However, demonstration of similar regulations between long-term epidermal growth factor (EGF) stimulation in pure human thyrocytes in culture and PTC supports the first interpretation (Hébrant *et al*., in revision).

Change in gene expression confirms previously identified tumor markers and is in accordance with the dedifferentiation status of theses tumors

Not surprisingly, a great number of known protein tumor markers were upregulated such as *KRT19*, *CITED1*, *LGALS3*, *FN1* (Prasad *et al*., 2005), *SPP1*, *TIMP1* (Hawthorn *et al*., 2004), *ECM1*, *MUC1*, *S100A4* (Zou *et al*., 2005). Furthermore, a decrease in *CRABP1* expression level has been proposed as a biomarker in PTC (Hawthorn *et al*., 2004).

As expected, genes involved in thyroid hormone synthesis were statistically downregulated (Table 1), in agreement with the conclusions of Huang *et al*. (2001). Such results are consistent with the dedifferentiation of carcinoma cells.

Genes involved in the growth factors signaling cascades are regulated

Activation of the EGF-, insulin-like growth factor (IGF)-, fibroblast growth factor (FGF)- and hepatocyte growth factor-mitogen-activated protein kinase (HGF-MAPK) signaling pathways are known to induce proliferation and could, consequently, participate in tumorigenesis. Regulation of genes involved in these pathways was therefore analysed separately.

While constitutive activation of the RAS/RAF/MEK/ERK signaling pathway is considered as the primary event in papillary thyroid carcinogenesis, no gene encoding for proteins involved in this pathway was

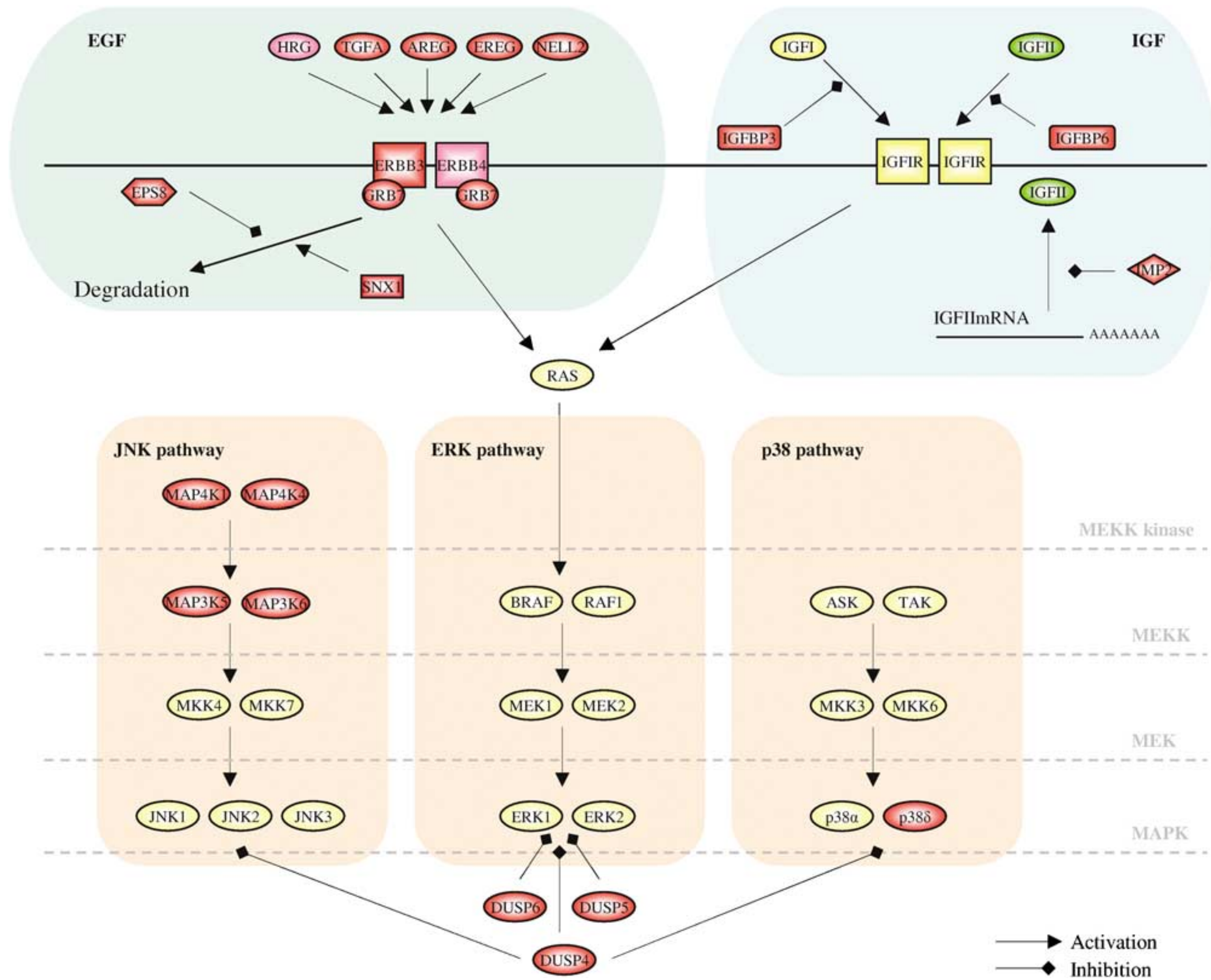

Figure 3 Simplified view of the EGF and IGF, and of the three main MAP kinase signaling pathways in PTC. See main text for explanations. Colors signification: dark red, overexpression in tumors in our data; bright red, overexpression in tumors from other studies; yellow, no regulation; green, underexpression in tumors. EGF, epidermal growth factor; IGF, insulin-like growth factor; MAP kinase, mitogen-activated protein kinase; PTC, papillary thyroid carcinoma.

dysregulated (Figure 3). However, rearranged RET or activating mutation of BRAF is sufficient to explain constitutive activation of this signaling pathway without overexpression of its components.

We also investigated different growth factor cascades that lead to the activation of this signaling pathway. EGF signaling has a central role in the pathogenesis and progression of different cancers (Normanno *et al*., 2006). However, the importance of this signaling pathway in PTC is controversial (Hoelting *et al*., 1994; Mitsiades *et al*., 2006). Our gene list shows a statistically significant alteration of the EGF signaling pathway (Table 1) and many genes involved in this cascade are overexpressed in our data (Figure 3): the EGF-like peptides (*TGFA*, *AREG*, *EREG* and *NELL2*), the *ERBB3* receptor and *GRB7*, a specific target of ERBB3 and ERBB4 (Fiddes *et al*., 1998). Moreover, heregulins, other EGF-related growth factors and ERBB4 were clearly upregulated in other PTC studies (Haugen *et al*., 1996; Fluge *et al*., 2000). On the other hand, *SNX1* (up) is involved in EGF receptor (EGFR) degradation (Kurten *et al*., 1996). However, *EPS8* (up) is shown to inhibit internalization of EGFR, and consequently, its degradation (Lanzetti *et al*., 2000). Thus, at least nine genes coding for proteins involved in EGF signaling cascade activation were upregulated for only one negative regulator, suggesting a positive balance in favor of activation of this pathway in PTC. As all ErbB ligands and receptors induce activation of the RAS/RAF/MEK/MAPK pathway (Normanno *et al*., 2006), this MAPK pathway is probably also activated in this way as well as by the constitutive activation of rearranged RET/PTC or mutated BRAF.

The IGF, FGF and HGF signaling pathways were also investigated in our gene list but no statistical results were found by DAVID. However, several inhibitors of the IGF signaling pathway were found upregulated and the IGF-II ligand was downregulated in our data, suggesting that this pathway is inhibited in PTC by regulation of its components (Figure 3).

Differential expression of genes involved in the c-jun N-terminal kinase (JNK) and p38 pathway was then investigated (Figure 3) and the activation of the JNK activity was found statistically significant by DAVID (Table 1). Four genes (*MAP4K1*, *MAP4K4*, *MAP3K5* and *MAP3K6*) encoding specific activators of the JNK but not of the ERK or p38 kinase pathways were upregulated in our gene list. This suggests that, beside ERK, the JNK pathway would be more activated in PTC compared to the adjacent tissue, as already suggested by an immunohistochemistry study (Shin *et al*., 2004). The upregulation of several cytokines in PTC (see above) could explain the activation of this cascade. On the other hand, three inhibitors of the MAPK pathways were highly overexpressed in PTC: *DUSP4*, *DUSP5* and *DUSP6* (Table 1). Interestingly, theses DUSP seem to show marked preference for ERK (Figure 3): DUSP5 and DUSP6 are highly specific for ERK, whereas DUSP4 inactivates ERK, p38 and JNK indifferently (Farooq and Zhou, 2004; Mandl *et al*., 2005). Thus, DUSP proteins could mitigate to some extent the activation of the ERK pathway, which could explain the positive but slow growth rate of the PTC.

Immediate early genes are downregulated
Upregulation of immediate early genes is an early step in the initiation of the cell cycle and some of them, when overexpressed in different models, induce cell proliferation (Milde-Langosch, 2005). 'Immediate early genes' is not a GO category; thus, we could not ascertain its status with DAVID. However, our results showed a general uncompensated repression of immediate early genes (0 up, 9 down; see Supplementary Table 4). This finding seems counterintuitive. Considering the low proliferation rate and proportion of cells in the cell cycle (for example, KI67 positive) in the tumor and the transient overexpression of these genes in the G_1 phase, an absence of upregulation would have been expected. However, the observed downregulation is more difficult to explain. A similar observation has been made in autonomous adenomas (Wattel *et al*., 2005).

Overexpression of many proteases and adhesion matrix proteins is consistent with the important remodeling in PTC
Aside from proliferation, other processes have to be altered to support tumor progression. Proteolytic degradation of the extracellular matrix (ECM) is an essential process for its remodeling, for migration and for metastasis (Skrzydlewska *et al*., 2005). In PCCL3 cells, the induction of genes coding for proteins with invasion properties is part of the gene response common to BRAF constitutive activation and RET/PTC3 rearrangement, and can thus be related to the constitutive activation of the RAS-MAPK pathway (Melillo *et al*., 2005). In contrast with follicular carcinoma, PTC, whatever the variant, has a strong stromal component. This connective tissue component present in papillae and fibrotic bands is part of the tumor and accompanies the tumor progression. This specific morphology clearly indicates a remodeling of the ECM, reflected in our data by a highly significant alteration of the ECM, a striking overexpression of proteases and protease inhibitors, and an alteration of adhesion matrix proteins, including fibronectin, collagens, laminins and elastins (Table 1 and Figure 4). The predominance of collagen overexpression (7 up, 1 down) is consistent with fibrosis of the tumors. Different groups of proteases leading to destruction of the ECM were upregulated in our data, including aspartyl proteases cathepsin D (*CTSD*), cysteine proteases (for example, cathepsin B (*CTSB*)), serine proteases (for example, *PLAU*) and metalloproteinases (*MMP1*, *MMP7* and *MMP11*) (Figure 4). A major process involved in carcinogenesis is the activation of several proteases by CTSD (Skrzydlewska *et al*., 2005).

CTSD can be autoactivated in an acidic environment, which is usually the case in the extracellular environment of tumors. It is able to degrade many ECM proteins, to inactivate cysteine protease inhibitors and to activate CTSB (Skrzydlewska *et al*., 2005) (Figure 4). Once activated, CTSB degrades the protein components of basement membranes and the interstitial connective matrix, including laminins, fibronectin and different types of collagens. It also activates metalloproteases and inactivates some MMP inhibitors (for example, TIMP1), facilitating progression of the tumor in the ECM (Skrzydlewska *et al*., 2005). Finally, CTSB is an activator of uPA (also called PLAU), whose expression was also induced in PTC, and which promotes invasiveness in follicular carcinoma cell lines (Sid *et al*., 2006). This protein is one of the two plasminogen activators that convert the zymogen plasminogen to the active serine protease plasmin. Plasmin is known to play a major role in the activation of several MMPs and induces ECM remodeling, facilitating cancer invasion and metastasis (Dano *et al*., 2005; Semov *et al*., 2005). Taken together, these results show that the mRNA of the main proteins (CTSD, CTSB and uPA) responsible for ECM remodeling, which can act directly by degrading ECM or indirectly by activating other proteases, are upregulated. This suggests that ECM remodeling occurs in PTC at least in part by these processes. The role of plasminogen activation seems to be particularly relevant in PTC because no inhibitors of this group of proteases are present in our gene list but also because it seems to be activated by another process: the activation of the second plasminogen activator (tPA) by upregulation of some S100 proteins and of annexin A2. Indeed, S100A4/ANXA2, S100A10/ANXA2 and probably S100A13/ANXA2 complexes increase the tPA-mediated plasmin production from plasminogen (Semov *et al*., 2005). These results thus suggest one additional mechanism for the extensive ECM remodeling occurring in PTC (Figure 4).

Our gene list is consistent with the tumor invasion mode of PTC
Different invasion mechanisms, separated in individual- and collective-cell migration modes, have been described in the literature (Friedl, 2004). These different strategies

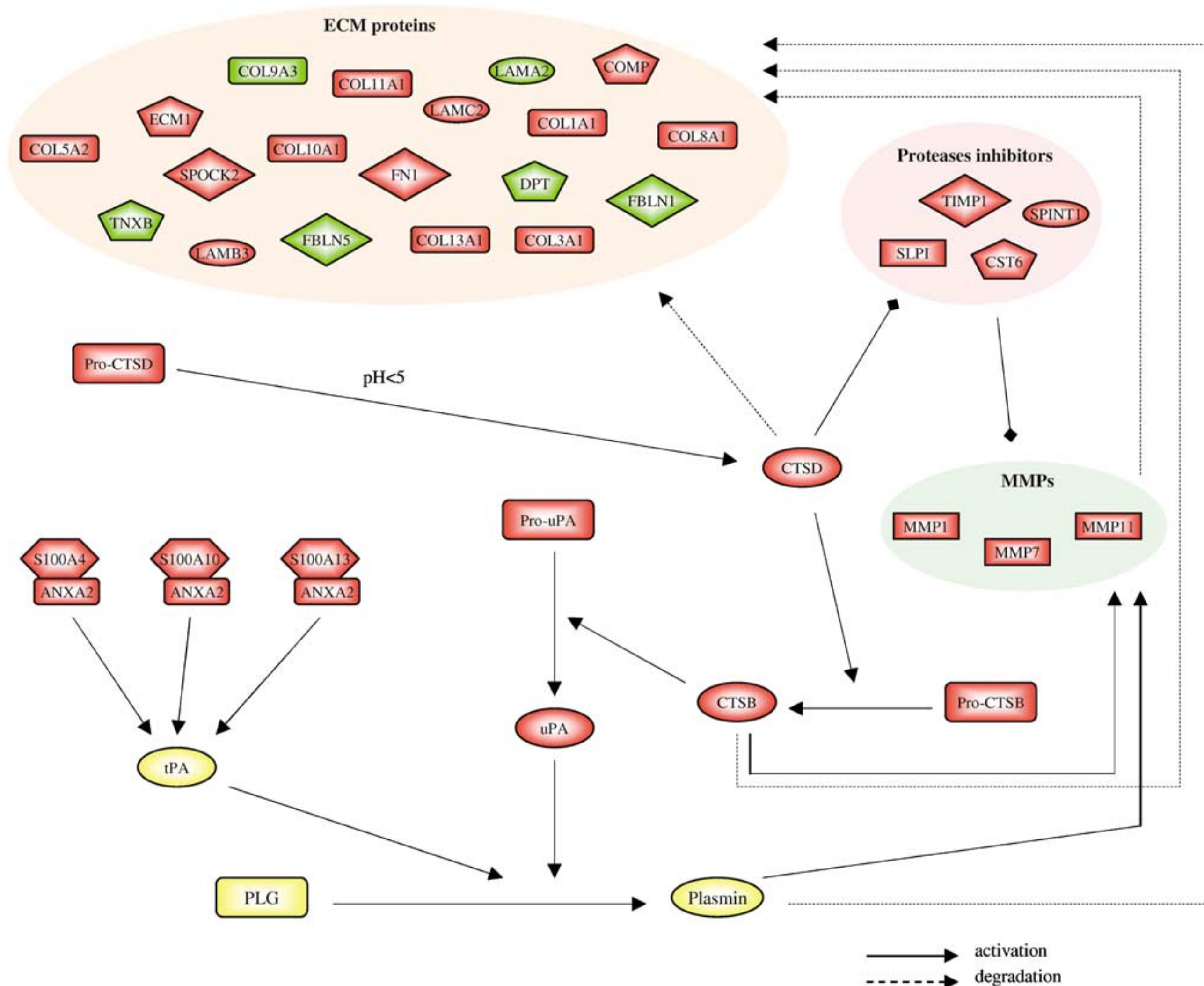

Figure 4 Simplified view of the ECM proteins degradation in PTC. See main text for explanations. Colors signification: red, overexpression in tumors; yellow, no regulation; green, underexpression in tumors. ECM, extracellular matrix; PTC, papillary thyroid carcinoma.

are determined by different molecular programmes with a higher expression of cell–cell adhesion proteins, proteases and integrins in collective- than in single-cell migration modes. Our gene list reveals a high proportion of upregulated genes coding for cell–cell adhesion proteins (17 up and 2 down), which was statistically significant according to DAVID (Table 1). They include different cadherins and claudins, which largely compensate the two downregulated genes, *CDH16* and *CLDN5* (see Supplementary Table 5). Moreover, five genes coding for integrin subunits were upregulated (see Supplementary Table 5) with a DAVID *P*-value slightly above 0.05 (Table 1) and overexpression of many proteases was observed (see above). These results are consistent with a predominant collective-cell migration mode of tumor cells, in clusters and sheets, observed by the pathologists (Figure 5; Friedl, 2004).

On the other hand, the epithelial to mesenchymal transition is usually considered as an important process leading to dissemination and metastasis spread in numerous cancers. The epithelial to mesenchymal transition markers (downregulation of CDH1 and cytokeratins and overexpression of vimentin) were not or inversely regulated in our gene list, suggesting that this process does not occur globally in the tumor, although it might happen locally, in highly invasive region, as recently proposed by Vasko *et al.* (2007). These results suggest an inverse relation between intercellular adhesion and distant metastasis ability. The fact that claudins are no longer overexpressed in the much more invasive dedifferentiated forms supports this hypothesis (Fluge *et al.*, 2006).

General conclusion

In conclusion, we compiled a reliable list of genes regulated in the majority of PTC from three large-scale independent microarray studies. This enabled us to better understand papillary thyroid carcinogenesis and

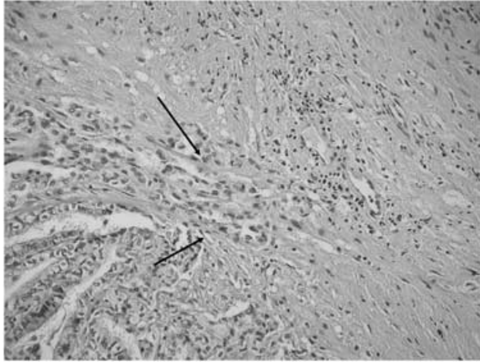

Figure 5 Hematoxylin eosin staining (× 200 magnification). The fibrous inflammatory interface between the tumor and the adjacent tissue is penetrated by tumor cells organized in sheets (showed by arrows), illustrating the collective-migration mode of tumor cells hypothesis applied to PTC progression. PTC, papillary thyroid carcinoma

to correlate many of these data to the biology and the histology of the tumor. Interestingly, alterations in the regulation of some signaling cascades described in this paper, such as activation of the EGF signaling pathway, are also found in mice models of PTC. In the future, these mice will enable us to confirm *in vivo* the biological role of candidate genes in papillary thyroid carcinogenesis.

Materials and methods

Tissue samples
Paired samples of tumoral and non-tumoral thyroid tissue counterparts were obtained from the Institute of Endocrinology and metabolism in Kiev, Ukraine via the Chernobyl Tissue Bank ($n = 12$, www.chernobyltissuebank.com) and from patients undergoing surgery for PTC at the A. Pare Hospital ($n = 14$; Boulogne, France). Clinical and histological information of these tissues are provided in Supplementary Table 6. Tumoral and tissue counterparts were immediately frozen in liquid nitrogen and stored at −80°C until use. Diagnoses were made by the Department of Pathology at the A. Pare Hospital in Boulogne or by the International Pathology Panel of the Chernobyl Tissue Bank samples. The protocol received approval from the Institutional ethics committees.

Microarray experiments
RNA purification, amplification, cDNA synthesis and labeling were performed as described in Supplementary Information. All the tumor/non-tumor tissue pairs ($n = 26$) were hybridized according to the manufacturer's protocol on Human 1 cDNA microarray (Agilent Technologies, Palo Alto, CA, USA) covering 8000 genes. The microarrays were scanned with a Genepix 4000B scanner. All details of microarray data preprocessing, normalization and detection of differentially expressed genes are described in Supplementary Information.

Real-time RT–PCR
Confirmation of the reliability of the gene list was performed by real-time RT–PCR (Eurogentec, Liege, Belgium) on 11–20 tumor/non-tumoral samples for 15 selected genes. The primers were designed with the Primer Express software (Applied Biosystems, Foster City, CA, USA) and are listed in Supplementary Table 7. All PCR efficiencies, obtained with four serial dilutions points (ranging from 20 ng to 200 pg), were above 90% and real-time RT–PCR was performed in duplicate for each gene. More details of the procedure are provided in Supplementary Information.

Normalization of real-time RT–PCR data
To identify the most stably expressed control genes for normalization in our thyroid cancer samples, several potential non-regulated genes were amplified by real-time RT–PCR in 20 tumor/non-tumor pairs of samples. *PBGD* (porphobilinogen deaminase) and *36B4* (ribosomal phosphoprotein acidic P0), two commonly used housekeeping genes, were tested and compared to *NEDD8* and *TTC1*, two of the most stable genes found in our microarray data set. Using the GeNorm software (Vandesompele *et al*., 2002), we identified the best combination of non-modulated genes to normalize our data. *PBGD* and *36B4* came out in first and second rank, respectively, while the best stability score was obtained by the combination of *NEDD8* and *TTC1* (*M*-value in GeNorm = 0.192). Because the association of these two genes gave a good stability score according to GeNorm, we used them as normalization genes for RT–PCR. The tumor/non-tumor expression ratio of a target gene was obtained by dividing their respective normalized quantities obtained by GeNorm (Vandesompele *et al*., 2002).

Acknowledgements

We thank Chantal Degraef for her excellent technical assistance. We acknowledge the confirmation of diagnosis provided by the International Pathology Panel of the Chernobyl Tissue Bank: Dr Alexandr Abrosimov, Professor Masahiro Ito, Professor Virginia LiVolsi, Professor Juan Rosai and Professor Sir Dillwyn Williams. This work was supported by Ministère de la Politique Scientifique (PAI), Action Concertée de la Communauté Française, Fonds National de la Recherche Scientifique, Fonds de la Recherche Scientifique Médicale, Télévie and Fondation Van Buuren. Laurent Delys is supported by Télévie, Vincent Detours by European Union's Marie Curie grant MEIF-CT-2003-501459, Frédérick Libert by the Fonds National de la Recherche Scientifique. The works of all authors were not cited in this paper due to a limitation of space. We apologize for it. Microarray data accession number: GSE3950 (GEO database, NCBI).

References

Alberti L, Carniti C, Miranda C, Roccato E, Pierotti MA. (2003). RET and NTRK1 proto-oncogenes in human diseases. *J Cell Physiol* **195**: 168–186.

Baris O, Mirebeau-Prunier D, Savagner F, Rodien P, Ballester B, Loriod B *et al*. (2005). Gene profiling reveals specific oncogenic mechanisms and signaling pathways in oncocytic and papillary thyroid carcinoma. *Oncogene* **24**: 4155–4161.

Borrello MG, Alberti L, Fischer A, Degl'innocenti D, Ferrario C, Gariboldi M *et al*. (2005). Induction of a proinflammatory

program in normal human thyrocytes by the RET/PTC1 oncogene. *Proc Natl Acad Sci USA* **102**: 14825–14830.
Ciampi R, Nikiforov YE. (2005). Alterations of the BRAF gene in thyroid tumors. *Endocr Pathol* **16**: 163–172.
Dano K, Behrendt N, Hoyer-Hansen G, Johnsen M, Lund LR, Ploug M *et al*. (2005). Plasminogen activation and cancer. *Thromb Haemost* **93**: 676–681.
de Visser KE, Eichten A, Coussens LM. (2006). Paradoxical roles of the immune system during cancer development. *Nat Rev Cancer* **6**: 24–37.
Dennis Jr G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC *et al*. (2003). DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol* **4**: 3.
Detours V, Wattel S, Venet D, Hutsebaut N, Bogdanova T, Tronko MD *et al*. (2005). Absence of a specific radiation signature in post-Chernobyl thyroid cancers. *Br J Cancer* **92**: 1545–1552.
Farooq A, Zhou MM. (2004). Structure and regulation of MAPK phosphatases. *Cell Signal* **16**: 769–779.
Fiddes RJ, Campbell DH, Janes PW, Sivertsen SP, Sasaki H, Wallasch C *et al*. (1998). Analysis of Grb7 recruitment by heregulin-activated erbB receptors reveals a novel target selectivity for erbB3. *J Biol Chem* **273**: 7717–7724.
Fluge O, Akslen LA, Haugen DR, Varhaug JE, Lillehaug JR. (2000). Expression of heregulins and associations with the ErbB family of tyrosine kinase receptors in papillary thyroid carcinomas. *Int J Cancer* **87**: 763–770.
Fluge O, Bruland O, Akslen LA, Lillehaug JR, Varhaug JE. (2006). Gene expression in poorly differentiated papillary thyroid carcinomas. *Thyroid* **16**: 161–175.
Friedl P. (2004). Prespecification and plasticity: shifting mechanisms of cell migration. *Curr Opin Cell Biol* **16**: QJ;14–23.
Gimm O. (2001). Thyroid cancer. *Cancer Lett* **163**: 143–156.
Giordano TJ, Kuick R, Thomas DG, Misek DE, Vinco M, Sanders D *et al*. (2005). Molecular classification of papillary thyroid carcinoma: distinct BRAF, RAS, and RET/PTC mutation-specific gene expression profiles discovered by DNA microarray analysis. *Oncogene* **24**: 6646–6656.
Haugen DR, Akslen LA, Varhaug JE, Lillehaug JR. (1996). Expression of c-erbB-3 and c-erbB-4 proteins in papillary thyroid carcinomas. *Cancer Res* **56**: 1184–1188.
Hawthorn L, Stein L, Varma R, Wiseman S, Loree T, Tan D. (2004). TIMP1 and SERPIN-A overexpression and TFF3 and CRABP1 underexpression as biomarkers for papillary thyroid carcinoma. *Head Neck* **26**: 1069–1083.
Hoelting T, Siperstein AE, Clark OH, Duh QY. (1994). Epidermal growth factor enhances proliferation, migration, and invasion of follicular and papillary thyroid cancer *in vitro* and *in vivo*. *J Clin Endocrinol Metab* **79**: 401–408.
Hosack DA, Dennis Jr G, Sherman BT, Lane HC, Lempicki RA. (2003). Identifying biological themes within lists of genes with EASE. *Genome Biol* **4**: R70.1–R70.8.
Huang Y, Prasad M, Lemon WJ, Hampel H, Wright FA, Kornacker K *et al*. (2001). Gene expression in papillary thyroid carcinoma reveals highly consistent profiles. *Proc Natl Acad Sci USA* **98**: 15044–15049.
Jarzab B, Wiench M, Fujarewicz K, Simek K, Jarzab M, Oczko-Wojciechowska M *et al*. (2005). Gene expression profile of papillary thyroid cancer: sources of variability and diagnostic implications. *Cancer Res* **65**: 1587–1597.
Kondo T, Ezzat S, Asa SL. (2006). Pathogenetic mechanisms in thyroid follicular-cell neoplasia. *Nat Rev Cancer* **6**: 292–306.
Kurten RC, Cadena DL, Gill GN. (1996). Enhanced degradation of EGF receptors by a sorting nexin, SNX1. *Science* **272**: 1008–1010.
Lanzetti L, Rybin V, Malabarba MG, Christoforidis S, Scita G, Zerial M *et al*. (2000). The Eps8 protein coordinates EGF receptor signalling through Rac and trafficking through Rab5. *Nature* **408**: 374–377.
Livolsi VA, Albores-Saavedra J, Asa Sl, Baloch ZW, Baloch ZW, Baloch ZW *et al*. (2004). *Pathology and Genetics of Tumours of Endocrine Organs*. In: DeLellis RA, LLoyd RV, Heitz PU, Eng Ch (eds). Oxford University press: Oxford, pp 57–66.
Mandl M, Slack DN, Keyse SM. (2005). Specific inactivation and nuclear anchoring of extracellular signal-regulated kinase 2 by the inducible dual-specificity protein phosphatase DUSP5. *Mol Cell Biol* **25**: 1830–1845.
Melillo RM, Castellone MD, Guarino V, De Falco V, Cirafici AM, Salvatore G *et al*. (2005). The RET/PTC-RAS-BRAF linear signaling cascade mediates the motile and mitogenic phenotype of thyroid cancer cells. *J Clin Invest* **115**: 1068–1081.
Milde-Langosch K. (2005). The Fos family of transcription factors and their role in tumourigenesis. *Eur J Cancer* **41**: 2449–2461.
Mitsiades CS, Kotoula V, Poulaki V, Sozopoulos E, Negri J, Charalambous E *et al*. (2006). Epidermal growth factor receptor as a therapeutic target in human thyroid carcinoma: mutational and functional analysis. *J Clin Endocrinol Metab* **91**: 3662–3666.
Normanno N, De Luca A, Bianco C, Strizzi L, Mancino M, Maiello MR *et al*. (2006). Epidermal growth factor receptor (EGFR) signaling in cancer. *Gene* **366**: 2–16.
Pavlidis P, Li Q, Noble WS. (2003). The effect of replication on gene expression microarray experiments. *Bioinformatics* **19**: 1620–1627.
Powell Jr DJ, Eisenlohr LC, Rothstein JL. (2003). A thyroid tumor-specific antigen formed by the fusion of two self proteins. *J Immunol* **170**: 861–869.
Prasad ML, Pellegata NS, Huang Y, Nagaraja HN, de la CA, Kloos RT. (2005). Galectin-3, fibronectin-1, CITED-1, HBME1 and cytokeratin-19 immunohistochemistry is useful for the differential diagnosis of thyroid tumors. *Mod Pathol* **18**: 48–57.
Semov A, Moreno MJ, Onichtchenko A, Abulrob A, Ball M, Ekiel I *et al*. (2005). Metastasis-associated protein S100A4 induces angiogenesis through interaction with annexin II and accelerated plasmin formation. *J Biol Chem* **280**: 20833–20841.
Shin E, Hong SW, Kim SH, Yang WI. (2004). Expression of down stream molecules of RET (p-ERK, p-p38 MAPK, p-JNK and p-AKT) in papillary thyroid carcinomas. *Yonsei Med J* **45**: 306–313.
Sid B, Dedieu S, Delorme N, Sartelet H, Rath GM, Bellon G *et al*. (2006). Human thyroid carcinoma cell invasion is controlled by the low density lipoprotein receptor-related protein-mediated clearance of urokinase plasminogen activator. *Int J Biochem Cell Biol* **38**: 1729–1740.
Skrzydlewska E, Sulkowska M, Koda M, Sulkowski S. (2005). Proteolytic–antiproteolytic balance and its regulation in carcinogenesis. *World J Gastroenterol* **11**: 1251–1266.
Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA *et al*. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* **102**: 15545–15550.
Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A *et al*. (2002). Accurate normalization of

real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* **3**: RESEARCH0034.01–0034.112.

Vasko V, Espinosa AV, Scouten W, He H, Auer H, Liyanarachchi S *et al*. (2007). Gene expression and functional evidence of epithelial-to-mesenchymal transition in papillary thyroid carcinoma invasion. *Proc Natl Acad Sci USA* **104**: 2803–2808.

Wattel S, Mircescu H, Venet D, Burniat A, Franc B, Frank S *et al*. (2005). Gene expression in thyroid autonomous adenomas provides insight into their physiopathology. *Oncogene* **24**: 6902–6916.

Zou M, Al Baradie RS, Al Hindi H, Farid NR, Shi Y. (2005). S100A4 (Mts1) gene overexpression is associated with invasion and metastasis of papillary thyroid carcinoma. *Br J Cancer* **93**: 1277–1284.

Supplementary Information accompanies the paper on the Oncogene website (http://www.nature.com/onc).