Gene expression in papillary thyroid carcinoma reveals highly consistent profiles

Ying Huang*, Manju Prasad†, William J. Lemon*, Heather Hampel*, Fred A. Wright*, Karl Kornacker†, Virginia LiVolsi§, Wendy Frankel†, Richard T. Kloos§, Charis Eng*, Natalia S. Pellegata*, and Albert de la Chapelle*†

*Human Cancer Genetics Program, Comprehensive Cancer Center, †Department of Pathology, Divisions of §Sensory Biophysics and ¶Endocrinology and Nuclear Medicine, Ohio State University, Columbus, OH 43210; and ‡Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA 19104

Contributed by Albert de la Chapelle, October 15, 2001

Papillary thyroid carcinoma (PTC) is clinically heterogeneous. Apart from an association with ionizing radiation, the etiology and molecular biology of PTC is poorly understood. We used oligo-based DNA arrays to study the expression profiles of eight matched pairs of normal thyroid and PTC tissues. Additional PTC tumors and other tissues were studied by reverse transcriptase–PCR and immunohistochemistry. The PTCs showed concordant expression of many genes and distinct clustered profiles. Genes with increased expression in PTC included many encoding adhesion and extracellular matrix proteins. Expression was increased in 8/8 tumors for 24 genes and in 7/8 tumors for 22 genes. Among these genes were several previously known to be overexpressed in PTC, such as MET, LGALS3, KRT19, DPP4, MDK, TIMP1, and FN1. The numerous additional genes include CITED1, CHI3L1, ODZ1, N33, SFTP, and SCEL. Reverse transcriptase–PCR showed high expression of CITED1, CHI3L1, ODZ1, and SCEL in 6/6 additional PTCs. Immunohistochemical analysis detected CITED1 and SFTP in 49/52 and 39/52 PTCs, respectively, but not in follicular thyroid carcinoma and normal thyroid tissue. Genes underexpressed in PTC included tumor suppressors, thyroid function-related proteins, and fatty acid binding proteins. Expression was decreased in 7/8 tumors for eight genes and decreased in 6/8 tumors for 19 genes. We conclude that, despite its clinical heterogeneity, PTC is characterized by consistent and specific molecular changes. These findings reveal clues to the molecular pathways involved in PTC and may provide biomarkers for clinical use.

Approximately 19,500 new cases of thyroid carcinoma are diagnosed each year in the United States, and 1,300 patients die of the disease (1). Papillary thyroid carcinoma (PTC) is the most common type of thyroid malignancy, accounting for about 80% of all thyroid cancers in the United States (2). The biological behavior of PTC varies widely, from indolent microcarcinomas, growing slowly with little or no invasion, to invasive tumors that metastasize and can cause death.

Although the etiology of PTC is generally poorly understood, a strong association exists with exposure to ionizing radiation in a minority of cases (3). Evidence from regions heavily contaminated by downfall from the Chernobyl accident show 100- to 200-fold increases in PTC, mainly in children (4). Most cases of PTC are sporadic, but as many as 6% of patients have a family history of PTC (5). Among all cancers not displaying regular Mendelian inheritance, thyroid cancer has the highest relative risk (8.60) for first-degree relatives of probands (6). These facts suggest the involvement of specific genes, including tumor suppressor genes and predisposing genes. However, genes, signaling pathways, and other basic mechanisms are currently poorly defined (7). Another shortcoming is the paucity of diagnostic and prognostic biomarkers.

This study was undertaken as a step toward identifying previously uncharacterized molecular genetic mechanisms in PTC. We chose to first analyze the levels of mRNA for more than 12,000 transcripts by microarray hybridization. Results from eight PTC tumors were compared with normal thyroid tissue from the same eight individuals. For selected genes the changes in expression were confirmed by semiquantitative reverse transcriptase (RT)-PCR. Protein expression was studied by immunohistochemistry in tumors from the eight patients and in additional tumors. We detected differential expression of genes that were previously known to be altered in PTC, validating the feasibility of our experimental approach. Numerous novel genes were found to be differentially expressed, some of them in a high proportion of the PTCs. For ease of discussion, we use the terms overexpression and underexpression to describe increased and decreased expression levels in the PTCs.

Materials and Methods

Tissue Samples. Tissues were snap-frozen in liquid nitrogen and stored at −80°C until use. Specimens were chosen for study based on two criteria: (i) histological diagnosis of PTC, including its follicular variant, and (ii) sufficient tissue of high purity (greater than 90% neoplastic cells in tumor; no neoplastic cells in normal tissue). Clinical data on the eight patients whose tumor and matching normal tissue samples were used in the microarray analysis are shown in Table 2 (which is published as supporting information on the PNAS web site, www.pnas.org). Samples from an additional six patients were processed similarly and studied for RNA and protein expression, but not by array analysis; data on these patients also are shown in Table 2.

Tissue Processing and Preparation of RNA. Frozen 20-μm sections were collected in test tubes and homogenized in RNeasy lysis buffer by applying the lysate onto a QIAshredder column (Qiagen, Valencia, CA). Total RNA was prepared by using the RNeasy Mini Kit (Qiagen). The integrity of the RNA was assessed by denaturing agarose gel electrophoresis (visual presence of sharp 28S and 18S bands) and spectrophotometry.

Microarray Analysis. Microarray analysis was performed as described in detail at http://www.cancergenetics.med.ohio-state.edu/microarray. Briefly, cRNA was prepared from 8 μg of total RNA, hybridized to HG-U95A Affymetrix oligonucleotide arrays (containing more than 12,000 human genes), scanned, and analyzed according to Affymetrix (Santa Clara, CA) protocols. Scanned image files were visually inspected for artifacts and normalized by using GENECHIP 3.3 software (Affymetrix). Comparisons were made for each tumor versus its matched normal sample, with the normal sample as baseline by using GENECHIP 3.3. The fold change values, indicating the relative change in the expression levels between the experimental (PTC tumor) and normal samples, were determined.

Abbreviations: PTC, papillary thyroid carcinoma; RT, reverse transcriptase.

To whom reprint requests should be addressed at: Human Cancer Genetics Program, Comprehensive Cancer Center, Ohio State University, Room 646, 454 West 12th Avenue, Columbus, OH 43210. E-mail: delachapelle-1@medcrtr.osu.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.
baseline targets (normal thyroid tissue), were used to identify genes differentially expressed between these two conditions.

**Data Analysis.** A file containing genes showing at least 2-fold change was produced for each sample pair. Genes were considered to have altered expression levels in PTC versus normal thyroid when meeting the following three criteria: (i) 2-fold or greater change in expression level in more than five of the eight pairs, (ii) $P < 0.05$ from paired $t$ test, and (iii) the difference between the means of average difference of the tumors and normals greater than 500. Although these criteria were somewhat arbitrary, they served as an effective means to identify a small group of genes with consistent differential expression in tumors vs. normal. In addition, estimates of gene expression were obtained according to the method of Li and Wong (8). This approach explicitly considers variation in hybridization intensity caused by individual probes on the array, producing improved expression estimates. For the cluster analyses, to correct for correlation between tumor and normal within a matched pair, supplemental data sets were created in which the mean expression for each pair was subtracted from the expression for each sample (http://thinker.med.ohio-state.edu). CLUSTER and TREEVIEW were used to cluster and visualize the data by using the correlation metric and average linkage (9). Hierarchical clustering was performed by using 1,202 genes remaining after filtering the data with SD $\geq 80$. The results described in Fig. 3 were derived from perfect match (PM) probe intensities by one of us (K.K.) by using a two-stage regression method similar to the PM-only method described by Li and Wong (8). Briefly, for each probe set, the PM probe sensitivities were first estimated by a regression approach. The gene expression level was then estimated as the slope of the regression line relating probe intensities to probe sensitivities (a consequence of the Li–Wong model). We performed this regression, and the $P$ value for each expression estimate was obtained from the $t$ statistic used to test the null hypothesis of zero slope. The $P$ values were compared with the Bonferroni corrected significance threshold $0.05/12,000$ to identify genes that were clearly expressed. In addition, in identifying potential PTC biomarkers we used $-\log(P)$ value directly as an alternative indicator of gene expression. We have found that such an approach has advantages in that it is relatively insensitive to any scaling procedures that are performed to make arrays comparable. Further information on these procedures is available on request from K.K. (kornacker@osu.edu).

**Semiquantitative RT-PCR.** One microgram of total RNA was incubated with DNase I and reverse-transcribed with oligo(dT) by using the Superscript II RT-PCR system (Life Technologies, Grand Island, NY). One microliter of RT product was amplified with primer pairs specific for the genes under study with $\beta$-actin transcript. Conditions and primer sequences are available on request. Each RT-PCR product was loaded on 2% agarose gel containing 0.5 $\mu$g/ml ethidium bromide. The gel image was made and saved by CHEMIIMAGER 4000 (version 4.04) imaging system (Alpha Inno-
The complete data set is contained at our web site (http://thinker.med.ohio-state.edu).

We found several genes that participate in these processes to be overexpressed in tumors relative to normal samples. (Fig. 2). Among these were TPO, DIO1, DIO2, and SLC5A5 (also named NIS, encoding sodium iodide symporter). Several additional related genes showed suggestive differential expression, but did not meet the rigid criteria for inclusion in Fig. 1, including TSHR (TSH receptor, 2-fold reduction in 3/8 PTCs, t statistic P = 0.09), and SLC26A4 (also named pendrin, down in 4/8 PTCs, P = 0.006). The results are consistent with the fact that most malignant thyroid tumors are hypofunctioning in trapping iodine and producing thyroid hormone (10).

**Results and Discussion**

**Differential Gene Expression Shows Consistent Patterns.** Our primary means of identifying differentially expressed genes was based on consistent fold-change and statistical significance assessed by paired t test for equality of expression in each tumor/normal pair. At least 2-fold overexpression in more than five PTCs was seen for 143 genes whereas at least 2-fold underexpression in more than five PTCs was seen for 83 genes. We assume that the genes of greatest interest are those that show concordant behavior of strong differential expression in many PTCs. We highlight genes with greater than 2-fold differential expression in Fig. 1: 24 genes overexpressed in 8/8 tumors, and 27 genes underexpressed in 7/8 or 6/8 tumors. Table 3, which is published as supporting information on the PNAS web site, provides more detailed information on these genes. The complete data set is contained at our web site (http://thinker.med.ohio-state.edu).

**Genes Related to Specialized Thyroid Functions.** The normal thyroid follicular cell is highly differentiated, having specialized properties including the ability to respond to thyroid-stimulating hormone (TSH), trap iodine, synthesize thyroglobulin, and maintain follicular structure (7). We found several genes that participate in these processes to be underexpressed in the PTCs (Table 2). Among these were TPO, DIO1, DIO2, and SLC5A5 (also named NIS, encoding sodium iodide symporter). Several additional related genes showed suggestive differential expression, but did not meet the rigid criteria for inclusion in Fig. 1, including TSHR (TSH receptor, 2-fold reduction in 3/8 PTCs, t statistic P = 0.09), and SLC26A4 (also named pendrin, down in 4/8 PTCs, P = 0.006). The results are consistent with the fact that most malignant thyroid tumors are hypofunctioning in trapping iodine and producing thyroid hormone (10).

**Overexpressed Genes Previously Associated with PTC.** Among the overexpressed genes, many are known from previous studies to be involved in PTC. A prototype example is FN1 (fibronectin 1), which is highly overexpressed in PTC (11). Other examples include MET (12), DPP4 (dipeptidylpeptidase IV) (13), SERPINA1 (a 1-antitrypsin) (14), KRT19 (keratin 19) (15), and LGAL3 (galectin-3) (16), which were overexpressed in all eight PTC samples. Several additional known genes were suggestive but did not meet the stringent criteria for Fig. 1, including

**LAMB3** (laminin B3, up in 6/8 PTCs, P = 0.003) (17), **MDK** (midkine, up in 7/8 PTCs, P = 0.005) (18), **MUC1** (mucin 1, up in 6/8 PTCs, P = 0.07) (19), and **TIMP1** (tissue inhibitor of metalloproteinase 1 up in 7/8 PTCs, P = 0.001) (20).
findings provide support for the design of the study and the
discovery of further novel genes.

Another gene showing major involvement in thyroid cancer is
the RET protooncogene, encoding a receptor tyrosine kinase
(21, 22). However, RET is among a small group of 25 genes with
fluctuating representation on different versions of the HG-
U95A array (www.affymetrix.com), showing inconsistent results
in array vs. RT-PCR comparisons.

Overexpressed Genes Not Previously Associated with PTC. Many
genes that we found overexpressed in PTC have previously been
associated with other cancers, cell cycle, or mitogenic control.
Examples include CITED1 (Cbp/p300-interacting transactivator
1, also known as melanocyte-specific gene 1), which is involved
in pigmented melanoma cells (23); N33, which is associated with
homozygous deletion in metastatic prostate cancer (24); SFTP
(surfactant, pulmonary-associated protein B) in pulmonary ad-
encarcinomas (25); CHI3L1 (cartilage glycoprotein-39), which
has been suggested to play a role in tumor invasion in colorectal
cancer (26), and EPS8 (epidermal growth factor receptor kinase
substrate), which is a substrate of receptor tyrosine kinases and
enhances epidermal growth factor-dependent mitogenic signals
(27). Genes not previously associated with any neoplasia or
thyroid disease include ADORA1, SCEL, ODZ1, PROS1,
KIAA0037, CST6, SDC4, P4HA2, DUSP6, TSSC3 (all seen in 8/8
tumors), and numerous others expressed in fewer than eight
samples. It is possible that as a result of more detailed studies in
the future many of these genes will turn out to have important
functions in PTC. The findings suggest the existence of yet
unexplored fundamental molecular pathways characterizing this
malignancy.

Overexpression of Genes Encoding Cell Adhesion-Associated Mole-
cules Is a Feature of PTC. Among the genes showing at least 2-fold
overexpression in 8/8 cases, cell adhesion molecules account for
as many as 10 of 24 (Fig. 1). Other cell adhesion molecules such as
TIMP1 (tissue inhibitor of metalloproteinase 1), LAMB3
(laminin, β 3), MDK (midkine or neurite growth-promoting
factor 2), and MUC1 (mucin 1) also showed substantial overex-
pression in a smaller subset of PTCs (see full data set at
http://tinker.med.ohio-state.edu). In contrast, cell adhesion
molecules account for only 3% of genes on the array for which
annotation could be obtained. Thus, we suggest that functional
subgroups such as cell adhesion molecules are specifically and
concordantly involved in many PTCs. It may be relevant that
some epithelial adhesion molecules are known to be associated
with the papillary growth pattern and high proliferative capacity
(17). How they possibly interact with each other and whether
they contribute to cell invasion and metastasis needs further
investigation.

Underexpressed Genes. Two-fold or greater underexpression was
not seen for any gene in 8/8 tumors. There were eight genes
underexpressed in 7/8 and 19 genes underexpressed in 6/8
tumors. These genes are shown by functional category in Fig. 1.
Not unexpectedly, some underexpressed genes encode proteins
involved in thyroid hormone metabolism, namely thyroid peri-
oxidase, and type I and II iodothyronine deiodinase. Notably,
several genes have previously been noted to have tumor sup-
pressor functions. Of particular interest are GAS1, CDC2L1, and
BCL2. Another finding is that several genes involved in fatty-acid
binding, such as FABP4, CRABP1, and APOD, are underex-
pressed in PTCs. The remaining genes have diverse or unknown
functions.

Cluster Analysis Shows Similarity of Expression Patterns Among
Tumors. Alternative estimates of gene expression were also
generated by using the model based approach proposed by Li
and Wong (8), which accounts for probe-specific effects. Paired
tests and results on fold changes were broadly supportive of the
earlier results. CITED1, CHI3L1, and SFTP were again the top the
differential expressed genes. ETV5, which had not been
identified in the previous analyses, was detected as overex-
pressed in all of the eight PTCs by using Li–Wong expression
estimates. Using either Affymetrix or model-based expression
estimates to cluster the tissues, each PTC tends to group with its
matched normal tissue, suggestive of constitutional similarity
shared by each paired sample (data not shown). To eliminate the
effect of person-to-person variation, we subtracted from each
expression value the average of the two expression values of that
gene for the matched pair. After this correction, all of the PTCs
were clustered together, and likewise all normal thyroid tissues
clustered together (Fig. 2). Such concordant clustering is un-
likely under chance variation (P < 0.008, sign test). Fig. 2 shows
two distinct profiles of gene expression. The first profile (Fig. 2h)
was represented by a group of 150 genes that were highly expressed in PTC tissues and underexpressed in the normal thyroid. Three fibronectin probes on the array were clustered together, which is supportive of the technical reproducibility of the array. Also clustered in this group are genes that have been highlighted in Fig. 1, as expected. Several genes encoding cell adhesion molecules, such as MDK, TIMP1, SERPINA1, MFGE8, and FN1, cluster into the same group, suggesting a similarity in expression pathways. Other gene expression relationships are also suggested by this clustering. For example, EPS8, ETV5, DUSP6, and ANXA1 cluster together, suggesting that they may occupy common signal transduction pathways. The second profile (Fig. 2c) was represented by a group of 180 genes that were more intensely expressed in normal thyroid tissue than in tumor tissue. Genes embedded within this cluster include transcripts that are related to thyroid functions, such as PAX8, SLC26A4 (pendrin), DIO1, DIO2, TSHR, TPO, and TG. Genes involved in fatty acid metabolism, such as FABP4, CRABP1, APOD, and ASAH, are also in this cluster. Underexpression of ASAH in PTC is in agreement with previous evidence (28). Genes belonging to the metallothionein family, such as MT1A, MT1B, MT1G, MT1F, and MT3, cluster closely to one another. These genes, all located in 16q13, encode heavy metal-binding proteins and have protective roles against cellular damage induced by UV radiation, heavy toxic metals, or reactive oxygen species (29, 30). This finding should stimulate experiments designed to explore the molecular mechanisms and consequences of underexpression of metallothionein genes in thyroid malignancy.

**Candidate Tumor Markers of PTC.** Thyroid nodules are observed in 4–7% of adults (31). Differential diagnosis usually is based exclusively on histology. Although PTC has relatively distinctive morphological features, specific biomarkers should be useful tools, especially in atypical cases such as follicular variants of PTC (32). A main criterion for a clinically useful marker is low or absent expression in normal tissue and high expression in tumor tissue. Fibronectin has been proposed as a marker to diagnose PTC (33). However, expression of fibronectin was detected in fibroblasts in normal thyroid, inducing false positive results (34). To allow us to search for genes fulfilling these criteria we devised a method based on the P values from the perfect match-only analysis described above. Among the genes on the array, four shown in Fig. 3 exhibit the greatest distinction between the PTCs and normal controls, whereas the P values among the controls are all nonsignificant when applying the Bonferroni criterion as a “detection” level (equal to 5.4 on the $-\log_{10}$ scale of the figure). By comparison, the expression level of fibronectin (FN1) was high in the tumors, but above the detection level in normal thyroid.

**Corroboration of Gene Expression.** We validated the differential expression of nine genes identified in this study by semiquantitative RT-PCR on two of the eight pairs analyzed by arrays and on six additional PTC cases. For each gene the RT-PCR results agreed well with the microarray data (Fig. 4). To further illustrate the concordance of the array and RT-PCR results, the bands were quantified and the tumor-to-normal ratio of signal strength was calculated. When these ratios were depicted together with ratios calculated from the fold changes by Affymetrix software (Fig. 6, which is published as supporting information on the PNAS web site) the concordance was found to be good.

**Immunohistochemical Analysis of Protein Expression.** Although both expression array and RT-PCR analysis give an estimate of the amount of transcript, a correlation with protein levels does not automatically follow. We studied the protein products of two highly overexpressed genes by means of immunohistochemical staining and took advantage of tissue microarrays to glean a more global view of the involvement of these proteins in thyroid cancer. For the CITED1 gene product, the results suggested a remarkable PTC specificity in that the protein was found in 39/42 PTC, 0/6 follicular thyroid carcinoma, 0/1 anaplastic thyroid carcinoma, and 0/9 normal thyroid tissues (Fig. 5, Table 1). Moreover, CITED1 was overexpressed in all 10 PTC cases.

**Fig. 5.** CITED1 and SFTPB are overexpressed in PTC. (a) Tissue microarray used for immunohistochemical analysis stained with anti-CITED1 antibody. (b) Representative areas of tissue microarrays stained with anti-CITED1 (C1, C16, and C53) and anti-SFTPB antibody (S13, S49, and S51). Normal thyroid does not stain (C53, S51), whereas in classical PTC (C1, S13) and follicular variant of PTC (C16, S49) there is strong staining. (c) Magnified fields corresponding to b. CITED1 and SFTPB are absent in follicular cells of normal thyroid, but strong in malignant cells of PTC. Magnifications: (a) ×2; (b) ×20; (c) ×400.
that we studied by DNA microarray and RT-PCR analysis. *CITED1* was expressed in both cytoplasm and nucleus. When tested on a tissue array containing 20 normal human tissues and 19 tumor tissues, *CITED1* was not expressed in any tissues except normal breast epithelium and weakly in one infiltrating ductal carcinoma of the breast (data not shown). Surfactant protein B, encoded by *SFTPB*, was not expressed in the majority of normal and tumor tissues, except alveolar type 2 cells in normal lung and lung cancer. However, *SFTPB* was strongly expressed in 33/42 PTCs, but not in nine normal thyroids (Fig. 5, Table 1). Thus, both *CITED1* and *SFTPB* appear to be PTC-specific.

**Conclusion.** As shown in recent studies of various cancers, expression arrays can provide insights that were hard to obtain when single genes or pathways were studied in the past. We show that PTC is no exception. Although PTC is clinically heterogeneous, the global expression patterns showed remarkable consistency. We recognize that further studies may disclose significant outliers as only eight matched pairs were analyzed on gene arrays and only six additional PTCs were studied to validate the findings. Nevertheless, this study provides a wealth of data, some of which confirmed previous knowledge whereas others are novel. In this first analysis of the data we concentrated on novel overexpressed genes, many of which may eventually point out dysregulated pathways. Some of these already show promise as clinical markers, being significantly up-regulated in the great majority of more than 40 PTCs. We anticipate that these findings will stimulate research into the molecular biology and clinical behavior of PTC.

We thank Drs. David Schuller and William Farrar for allowing us to use specimens they provided and Dr. Toshi Shioda for providing rabbit polyclonal anti-CITED1 antibody. This work was supported by Grant CA 16058 from the National Institutes of Health.
Table 2. Clinical features and expression of thyroid function-related genes in 14 cases of PTC

<table>
<thead>
<tr>
<th>Case no. (tissue no.)</th>
<th>Age</th>
<th>Sex</th>
<th>Follicular variant</th>
<th>Tumor size (cm)</th>
<th>Lymph node metastases</th>
<th>Extra-thyroidal extension</th>
<th>SLC26A4</th>
<th>TSHR</th>
<th>TPO</th>
<th>DIO1</th>
<th>DIO2</th>
<th>NIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (T1/N1)*</td>
<td>45</td>
<td>F</td>
<td>No</td>
<td>1.1</td>
<td>No</td>
<td>No</td>
<td>-31.3</td>
<td>-3.2</td>
<td>-10.8</td>
<td>-97.7</td>
<td>-3.3</td>
<td>Down</td>
</tr>
<tr>
<td>2 (T2/N2)*</td>
<td>60</td>
<td>M</td>
<td>No</td>
<td>3</td>
<td>Yes</td>
<td>Yes</td>
<td>-27.4</td>
<td>-4.6</td>
<td>-2.2</td>
<td></td>
<td></td>
<td>Down</td>
</tr>
<tr>
<td>3 (T3/N3)*</td>
<td>62</td>
<td>M</td>
<td>No</td>
<td>3</td>
<td>Yes</td>
<td>No</td>
<td>-100.2</td>
<td>-4.1</td>
<td>-103.1</td>
<td>-10</td>
<td>-4</td>
<td>Down</td>
</tr>
<tr>
<td>4 (T4/N4)*</td>
<td>48</td>
<td>F</td>
<td>Yes</td>
<td>2</td>
<td>Yes</td>
<td>No</td>
<td>-3.1</td>
<td>-2.5</td>
<td>-3.1</td>
<td></td>
<td></td>
<td>Down</td>
</tr>
<tr>
<td>5 (T5/N5)*</td>
<td>23</td>
<td>F</td>
<td>No</td>
<td>2</td>
<td>Yes</td>
<td>No</td>
<td>-26.1</td>
<td></td>
<td>-48.3</td>
<td>-6.9</td>
<td>-2.9</td>
<td>Down</td>
</tr>
<tr>
<td>6 (T6/N6)*</td>
<td>68</td>
<td>F</td>
<td>Yes</td>
<td>1</td>
<td>No</td>
<td>No</td>
<td>-2.8</td>
<td>-2.7</td>
<td>-2.4</td>
<td>-2.5</td>
<td>-2.4</td>
<td>Down</td>
</tr>
<tr>
<td>7 (T7/N7)*</td>
<td>60</td>
<td>F</td>
<td>No</td>
<td>0.9</td>
<td>No</td>
<td>No</td>
<td>-2.7</td>
<td>-2.7</td>
<td></td>
<td></td>
<td></td>
<td>Down</td>
</tr>
<tr>
<td>8 (T8/N8)*</td>
<td>51</td>
<td>F</td>
<td>No</td>
<td>4</td>
<td>Yes</td>
<td>No</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>Down</td>
</tr>
<tr>
<td>9 (T9/N9)</td>
<td>41</td>
<td>F</td>
<td>Yes</td>
<td>1.5</td>
<td>No</td>
<td>No</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>Down</td>
</tr>
<tr>
<td>10 (T10/N1)</td>
<td>48</td>
<td>M</td>
<td>No</td>
<td>3.8</td>
<td>No</td>
<td>Yes</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>Down</td>
</tr>
<tr>
<td>11 (T11/N1)</td>
<td>30</td>
<td>F</td>
<td>No</td>
<td>2.4</td>
<td>Yes</td>
<td>No</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>Down</td>
</tr>
<tr>
<td>12 (T12/N1)</td>
<td>30</td>
<td>F</td>
<td>No</td>
<td>0.9</td>
<td>Yes</td>
<td>Yes</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>Down</td>
</tr>
<tr>
<td>13 (T13/N1)</td>
<td>31</td>
<td>M</td>
<td>No</td>
<td>1.8</td>
<td>Yes</td>
<td>No</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>Down</td>
</tr>
<tr>
<td>15 (T15)</td>
<td>21</td>
<td>F</td>
<td>Yes</td>
<td>4.5</td>
<td>Yes</td>
<td>Yes</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>Down</td>
</tr>
</tbody>
</table>

SLC26A4, encoding pendrin; TSHR, encoding thyroid-stimulating hormone receptor; TPO, encoding thyroid peroxidase; DIO1 and DIO2 encoding type I and type II iodothyronine deiodinase, respectively; NIS, also named SLC5A5, encoding sodium iodide symporter; Negative number is fold change indicating decrease in expression in PTC in comparison with normal tissue. Information for NIS is derived from RT-PCR, so only “down” is shown. Empty space means the change is not significant. *: samples used in DNA array. ND: not determined.
Table 3. Concordantly overexpressed and underexpressed genes in PTC

<table>
<thead>
<tr>
<th>Fold change</th>
<th>Gene symbol</th>
<th>Gene name</th>
<th>Acc. no.</th>
<th>Location</th>
<th>Frequency</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.3 (6.1, 17.5)</td>
<td>FN1*</td>
<td>Fibronectin 1</td>
<td>X02761</td>
<td>2p34</td>
<td>8</td>
<td>0.0002</td>
</tr>
<tr>
<td>13.5 (7.6, 23.9)</td>
<td>SERPINA1</td>
<td>Alpha 1-antitrypsin</td>
<td>X01683</td>
<td>14q32.1</td>
<td>8</td>
<td>0.002</td>
</tr>
<tr>
<td>3.8 (3.1, 4.6)</td>
<td>LGALS3</td>
<td>Galectin-3</td>
<td>Y08374</td>
<td>2q34</td>
<td>8</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>27.7 (13.1, 58.7)</td>
<td>P4HA2</td>
<td>Prolyl 4-hydroxylase alpha (II) subunit</td>
<td>U00441</td>
<td>3q1</td>
<td>8</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>12.5 (6.9, 22.6)</td>
<td>ST6</td>
<td>Cystatin E/M</td>
<td>Y08374</td>
<td>14q32.1</td>
<td>8</td>
<td>0.002</td>
</tr>
<tr>
<td>9.0 (5.9, 13.9)</td>
<td>CST6</td>
<td>Cystatin E/M</td>
<td>W06521</td>
<td>1q13</td>
<td>8</td>
<td>0.0003</td>
</tr>
<tr>
<td>4.5 (3.4, 6.1)</td>
<td>CTSC4</td>
<td>Liver dipeptidyl peptidase IV</td>
<td>X60708</td>
<td>2q34</td>
<td>8</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>6.3 (4.3, 9.2)</td>
<td>KRT19</td>
<td>Keratin 19</td>
<td>Y00503</td>
<td>1q34</td>
<td>8</td>
<td>0.01</td>
</tr>
<tr>
<td>4.7 (3.2, 7.0)</td>
<td>SCEL*</td>
<td>Sciellin</td>
<td>AF045941</td>
<td>14q32.1</td>
<td>8</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>3.8 (3.0, 4.7)</td>
<td>EPS8</td>
<td>Epidermal growth factor receptor kinase substrate 8</td>
<td>U12535</td>
<td>12q23-q24</td>
<td>8</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>4.6 (3.7, 5.8)</td>
<td>DUSP6</td>
<td>Dual specificity phosphatase 6</td>
<td>AB013382</td>
<td>12q22-q23</td>
<td>8</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>26.9 (13.5, 53.7)</td>
<td>CITED1*</td>
<td>Cbp/p300-interacting transactivator 1</td>
<td>U65092</td>
<td>Xq13.1</td>
<td>8</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>14.5 (7.2, 29.4)</td>
<td>RIL</td>
<td>353kDa LIM domain protein</td>
<td>X93510</td>
<td>5q31</td>
<td>8</td>
<td>0.003</td>
</tr>
<tr>
<td>5.6 (4.9, 8.9)</td>
<td>MET</td>
<td>Met protooncogene</td>
<td>L22214</td>
<td>1q32.1</td>
<td>8</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>1.9 (1.6, 2.2)</td>
<td>ADORA1*</td>
<td>Adenosine A1 receptor</td>
<td>AF001294</td>
<td>11p15.5</td>
<td>8</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>3.5 (2.9, 4.2)</td>
<td>SDC4</td>
<td>Ryudocan core protein</td>
<td>D79206</td>
<td>20q12-q13</td>
<td>8</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>8.2 (5.4, 13.9)</td>
<td>CST6</td>
<td>Cystatin E/M</td>
<td>W68521</td>
<td>14q32.1</td>
<td>8</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>4.5 (3.4, 6.1)</td>
<td>CTSC4</td>
<td>Liver dipeptidyl peptidase IV</td>
<td>X60708</td>
<td>2q34</td>
<td>8</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

These genes were selected based on fold change, paired t test, and absolute expression values, and ranked according to the number of PTC samples in which each gene was differentially expressed. The genes are arranged into functional categories based on published evidence. The geometric mean fold-change of each gene in PTCs relative to normal tissues and 95% confidence interval, the P value by paired t test, GenBank accession number, chromosomal location, and number of sample pairs (of 8) in which the change occurred are shown. Negative fold-change means decrease in expression in PTCs in comparison to normal tissues. Genes marked by * are those confirmed by semiquantitative RT-PCR.