

# Perspectives and Limitations of Microarray-Based Gene Expression Profiling of Thyroid Tumors

Markus Eszlinger, Knut Krohn, Aleksandra Kukulska, Barbara Jarzab, and Ralf Paschke

III. Medical Department (M.E., K.K., R.P.), University of Leipzig, D-04103 Leipzig, Germany; Interdisciplinary Center for Clinical Research Leipzig (K.K.), D-04103 Leipzig, Germany; and Department of Nuclear Medicine and Endocrine Oncology (A.K., B.J.), Maria Skłodowska-Curie Memorial Cancer Center and Institute of Oncology, Gliwice Branch, 44-100 Gliwice, Poland

Microarray technology has become a powerful tool to analyze the gene expression of tens of thousands of genes simultaneously. Microarray-based gene expression profiles are available for malignant thyroid tumors (*i.e.*, follicular thyroid carcinoma, and papillary thyroid carcinoma), and for benign thyroid tumors (such as autonomously functioning thyroid nodules and cold thyroid nodules). In general, the two main foci of microarray investigations are improved understanding of the pathophysiology/molecular etiology of thyroid neoplasia and the detection of genetic markers that could improve the differential diagnosis of thyroid tumors. Their results revealed new features, not known from one-gene studies. Simultaneously, the increasing number of microarray analyses of different thyroid pathologies raises the demand to efficiently compare the data. However, the use of different microarray platforms complicates cross-analysis. In addition, there are other important differences between these studies:

1) some studies use intraindividual comparisons, whereas other studies perform interindividual comparisons; 2) the reference tissue is defined as strictly nonnodular healthy tissue or also contains benign lesions such as goiter, follicular adenoma, and hyperplastic nodules in some studies; and 3) the widely used Affymetrix GeneChip platform comprises several GeneChip generations that are only partially compatible. Moreover, the different studies are characterized by strong differences in data analysis methods, which vary from simple empiric filters to sophisticated statistic algorithms. Therefore, this review summarizes and compares the different published reports in the context of their study design. It also illustrates perspectives and solutions for data set integration and meta-analysis, as well as the possibilities to combine array analysis with other genetic approaches. (*Endocrine Reviews* 28: 322–338, 2007)

- I. Introduction
  - A. Potential of the method
  - B. Methodological evolution
- II. New Perspectives Generated by Gene Expression Profiling of Benign and Malignant Thyroid Tumors
  - A. Papillary thyroid carcinoma (PTC)
  - B. Follicular thyroid carcinoma (FTC)
  - C. Autonomously functioning thyroid nodules (AFTNs)
  - D. Cold thyroid nodules (CTNs)
  - E. Clues for the comparison and differential diagnosis of thyroid nodules, based on the microarray data
- III. First Results, Possibilities, and Solutions for Data Set Integration and Meta-Analysis
  - A. Study design
  - B. Types and generations of microarrays (one-color systems *vs.* two-color systems) and GeneChip generations
  - C. Selection of reference tissue and intra- *vs.* interindividual comparisons

- IV. Future Developments
  - A. Diagnostic implications
  - B. Genome analysis (SNP arrays, arrayCGH)

## I. Introduction

### A. Potential of the method

ALTHOUGH THE BASIC concept of “microarrays” was presented in the early 1990s, about 10 yr elapsed before microarray technology became a widely used tool to analyze gene expression of tens of thousands of genes simultaneously on high-density microarrays in a single experiment. Such gene expression data allow us to identify genes that are expressed in a given cell type at a particular time and under a particular condition, to identify key players or target genes in signaling pathways, to recognize new targets of drugs, to find molecular markers in disease diagnosis, or to investigate the genetic background of an individual’s response to environmental factors (1).

### B. Methodological evolution

As the sequence of information for the whole human genome became available, life scientists were challenged with the task of measuring the expression levels on a global scale. For this purpose, cDNA library clones were spotted on large membrane sheets to hybridize radiolabeled cDNA pools (2) generated from total or purified mRNA. This made it possible to measure the signal intensity of hundreds of transcripts. These ancestors of microarrays called “macroarrays”

First Published Online March 12, 2007

Abbreviations: AFTN, Autonomously functioning thyroid nodule; CGH, comparative genomic hybridization; CTN, cold thyroid nodule; FDR, false discovery rate; FNA, fine-needle aspiration; FNAB, FNA biopsy; FNAC, FNA cytology; FTC, follicular thyroid carcinoma; PKC, protein kinase C; PTC, papillary thyroid carcinoma; RFR, recursive feature replacement; RGS, regulator of G protein signaling; SNP, single nucleotide polymorphism; ST, surrounding tissues; SVM, support vector machine; TSHR, TSH receptor.

*Endocrine Reviews* is published by The Endocrine Society (<http://www.endo-society.org>), the foremost professional society serving the endocrine community.

were characterized by several difficulties; *e.g.*, they were made of porous material, they were difficult to handle, and they often required radioactive labeling techniques. The elimination of these technical drawbacks by the introduction of new carrier materials (*i.e.*, glass slides) and fluorescence-based signal detection in 1995, when the first real cDNA microarray was established (3), formed the basis for a broad application. Subsequently, new developments in automation, robotic spotting, photolithography, fluorescence-based detection, and bioinformatics led to a rapid growth in microarray usage. According to the length of spotted DNA fragments, all RNA expression arrays can be subdivided into two groups: 1) cDNA microarrays that use approximately 200- to 500-bp double-stranded DNA fragments, which are normally produced by PCR; and 2) oligonucleotide microarrays that use 25–70 bp of single stranded DNA. Both cDNA and oligonucleotide fragments are chemically attached to the glass slides. These arrays are traditionally used with a two-color detection system; *i.e.*, RNA from sample and control is labeled with fluorescent Cy3 and Cy5 dyes and hybridized to the same array. Therefore, these two-color arrays produce a ratio indicating the differential expression between the sample and the control. In contrast to these microarrays that are characterized by a deposition of cDNA fragments or oligonucleotides on the slides, Fodor and co-workers (Affymetrix, Santa Clara, CA) (4–7) pioneered the development of a new *in situ* synthesis technology that combines photolithography known from the semiconductor industry and chemical DNA synthesis. This technology enabled a further miniaturization of the assay and the manufacturing of high-density oligonucleotide microarrays (GeneChips): the current version of Affymetrix GeneChips (Human Genome U133 Plus 2.0) provides a comprehensive coverage of the transcribed human genome on a single array covering over 47,000 transcripts represented by more than 1 million distinct oligonucleotide entities that cover a 11- $\mu\text{m}$  square on the array (called feature size). Moreover, in contrast to the two-color arrays, Affymetrix GeneChips use a standardized biotin-labeling protocol and produce an intensity signal allowing absolute quantification.

The power of microarray studies depends not only on the quality of array design and production but also on the statistical and bioinformatic approaches used to analyze the data. Indeed, the application of different mathematical algorithms can influence the outcome of microarray data analysis enormously, *e.g.*, by different statistical powers to detect significant differentially expressed genes. Early studies and studies employing radioactively labeled macroarrays often used housekeeping genes for data normalization (8) or simple algorithms such as z-score transformation (9, 10). However, normalization to classical housekeeping genes such as glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) and  $\beta$ -actin (*ACTB*) is of limited value and requires special attention because it cannot be excluded that these genes are regulated [*e.g.*, whereas no regulation of these genes has been observed in AFTNs, they do show differential expression in TSH-stimulated primary thyroid epithelial cells (11)]. As an alternative, various preprocessing techniques have been developed: *e.g.*, MAS5.0 (12), dChip (13), Robust Multichip Average (RMA) (14, 15), and GC-RMA (16). However, so far

a clear recommendation for the best performing algorithm cannot be given (17). Similar to the high number of preprocessing algorithms, numerous methods of data analysis have been published. They can be classified as methods to detect differentially expressed genes and gene sets in a supervised manner (*e.g.*, empirical filtering, statistic algorithms such as *t* test, F test), unsupervised cluster analysis (*e.g.*, hierarchical clustering, self-organizing maps, k-mean clustering), and supervised sample classification [*e.g.*, Recursive Feature Replacement (RFR) (18, 19), Prediction Analysis of Microarrays PAM (20)].

Because the introduction of microarrays has offered the possibility to investigate a multitude of parameters (*i.e.*, genes) in a small number of samples (*i.e.*, microarrays), this possibility introduced in parallel the most challenging problem of microarray analysis, which is the so-called “problem of multiple comparisons.” Therefore, in up-to-date microarray studies, a correction for multiple comparisons [*e.g.*, false discovery rate (FDR), Westfall-Young step-down permutation correction (21), and significance analysis of microarrays (22)] should be mandatory.

Hierarchical cluster analysis has become the most popular and most frequently used multivariate technique to analyze microarray data that results in a complete tree with leaves as individual patterns (genes or experiments) and the root as the convergence point of all branches. However, given enough genes, the genes will always cluster. Therefore, there is only minor scientific value in the fact that there are genes that behave in a similar way. Instead, the interpretation of a common clustering may be crucial. However, sometimes it seems that authors feel the urge to present their microarray results in the form of a cluster diagram, although such clusters are not always meaningful.

Because the conventional diagnosis of cancer is based on the morphological appearance of stained tissues whose analysis requires highly trained pathologists, the introduction of microarrays offered the hope that classification and prediction of cancer by means of their gene expression profiles could be more objective and accurate. Nevertheless, the problem of classification by microarrays is not simple; the genes that make it possible to predict the classes have to be identified from a large number of genes (in a relatively small number of samples). Moreover, it is important to identify which genes contribute most to the classification. For this purpose, several new computational methods have been developed. In addition to algorithms suggested by Golub *et al.* (23) and Hedenfalk *et al.* (24), a simple and well-performing approach based on nearest shrunken centroids has been proposed (20). At the moment, the reference method for classification problems in microarray studies are support vector machine (SVM) algorithms, and new promising supervised gene selection algorithms based on the SVM technique (25) [such as recursive feature elimination (26) and RFR (18, 19)] have been introduced.

An increasing number of studies have used microarrays in the field of thyroidology. These studies investigate (patho)physiological questions or aim at elucidating the tumor’s molecular etiology (Table 1), whereas others look primarily for genetic markers that could improve the differential diagnosis of thyroid tumors. Both platforms, spotted cDNA

TABLE 1. Contributions of microarray studies to unravel the molecular etiology or (patho)physiology of thyroid tumors

Tumor entity	Findings
PTC	Strong relationship between mutational status ( <i>BRAF</i> , <i>RET/PTC</i> , <i>RAS</i> ), gene expression, and morphology in PTC (46) <i>RET/PTC3</i> -, <i>HRAS</i> -, and <i>BRAF</i> -transfected PC Cl3 cells show a widely overlapping set of differentially expressed genes (58) <i>RET/PTC</i> - <i>RAS</i> - <i>BRAF</i> linear signaling cascade mediates the motile and mitogenic phenotype of thyroid cancer cells (58)
FTC	Induction of a proinflammatory program in normal human thyrocytes by the <i>RET/PTC1</i> oncogene (56) FTC with a <i>PAX8-PPAR<math>\gamma</math></i> rearrangement form a distinct biological entity within the FTC (36)
AFTNs; TSH-stimulated primary cultures	Suppression of negative feedback modulators of the cAMP cascade in AFTNs (11) Inactivation of the TGF $\beta$ -signaling pathway in AFTNs (52) Relevance of sialylation for TSHR signaling (52, 84) Long-term TSH-stimulated primary cultures as model system for AFTNs (11, 51)
CTNs	Expression pattern of cell cycle-associated genes supports increased proliferation of CTNs (54) Major relevance of G $_q$ -PKC-signaling in CTNs (54)

and oligonucleotide microarrays with a two-color or radioactive detection, and *in situ* photolithographically synthesized oligonucleotide arrays with a one-color detection, have been used to analyze the gene expression profiles of malignant thyroid tumors [*i.e.*, follicular thyroid carcinoma (FTC) (27–37) and papillary thyroid carcinoma (PTC) (31–33, 37–50)], and benign thyroid tumors [such as autonomously functioning thyroid nodules (AFTNs) (8, 11, 48, 51–53) and cold thyroid nodules (CTNs) (8, 48, 51, 54)].

## II. New Perspectives Generated by Gene Expression Profiling of Benign and Malignant Thyroid Tumors

### A. Papillary thyroid carcinoma (PTC)

The gene expression pattern of PTC, the most common type of thyroid malignancy, which accounts for approximately 80% of all thyroid cancers in the United States (55), is the thyroid tumor most extensively investigated by array studies (31–33, 37–40, 42–50, 56–58). The first array investigation of PTC was performed by Huang *et al.* (38) in 2001 using U95A Affymetrix GeneChips comprising more than 12,000 transcripts. In their study of eight PTC tumors, which were compared with normal surrounding tissue from the same eight individuals, the authors specified 50 genes with the most distinct gene expression changes. They confirmed a decreased expression of several thyroid-specific genes [*e.g.*, *TPO* (thyroid peroxidase), *DIO1* (type I iodothyronine deiodinase), *DIO2* (type II iodothyronine deiodinase), and *SLC5A5* (sodium iodide symporter)], which reflected the fact that most malignant thyroid tumors are hypofunctioning (59), and observed an increased expression of several genes already known for differential expression in PTC [*e.g.*, *FN1* (fibronectin 1), *MET* (met proto-oncogene), *DPP4* (dipeptidyl-peptidase 4), *SERPINA1* (serpin peptidase inhibitor, clade A, member 1), *KRT19* (keratin 19), and *LGALS3* (galactin 3)], which supported the reliability of their analyses. Moreover, they identified a number of additional PTC-specific genes, many of them associated with cell cycle or mitogenic control [*e.g.*, *CITED1* (Cbp/p300-interacting transactivator 1) and *DUSP6* (dual specificity phosphatase 6)], and some of them previously found in other cancers. This study especially stressed the overexpression of genes encoding cell adhesion-associated molecules [besides *FN1*, *DPP4*, *SER-*

*PINA1*, and *LGALS3*, also *SFTPB* (surfactant, pulmonary-associated protein B), *TIMP1* (TIMP metalloproteinase inhibitor 1), or *MUC1* (mucin 1, cell surface associated)]. In subsequent studies, many of the genes specified in this study (*e.g.*, *FN1*, *TIMP*, *MET*, *LGALS3*, and *DUSP6*) were confirmed to distinguish between PTC and normal or benign thyroid tissue independent of the microarray platform and the analysis algorithms used (33, 40, 44). Also, the findings of Jarzab *et al.* (44) reflected the importance of adhesion genes in PTC (38). These genes constituted the most frequent gene ontology class in their data set (17% of the genes that are crucial for tumor/normal thyroid difference). Thus, the involvement of adhesion-related genes appears a characteristic feature of PTC, which is not so distinct in many other cancers.

For the adhesion-controlling PTC genes, it is well known that they participate in both invasion and metastasis and are involved in inflammation. Borrello *et al.* (56), who investigated the specific gene expression patterns of primary human thyrocytes transfected with the *RET/PTC1* oncogene, could show that this oncogene activates the expression of a large set of genes involved in inflammation and tumor invasion, such as chemokines, matrix-degrading enzymes, and adhesion molecules. Interestingly, they also found this expression pattern in specimens of PTC harboring the *RET/PTC* rearrangement and pT4N1 presentation, which validates the *in vivo* relevance of their *in vitro* results. Therefore, and because PTCs are often associated with chronic inflammatory thyroiditis, the authors concluded that the inflammation-associated genes activated by *RET/PTC1* most likely contribute to tumor progression and locoregional metastasis (56). This conclusion remains in some conflict with clinical data showing better outcome in PTC cases with concomitant lymphocytic infiltration (60). In the PTC gene expression profile, a very distinct expression pattern of immune response genes was observed, which follows in terms of intensity the pattern that is characteristic for the difference between tumor and normal tissue (44).

The molecular background of PTC was comprehensively investigated in a recent study of Melillo *et al.* (58). Using a broad methodology, they showed that the oncogenic proteins (*i.e.*, *RAS*, *BRAF*, *RET/PTC* rearrangements) involved in the initiation of PTC signal within one cascade, namely the *RAS*-*BRAF*-*MAPK*-pathway. Nearly two thirds of the targets that are under the control of *RET/PTC* were shown to be

regulated by the RAS-BRAF-MAPK pathway. This finding is supported by the gene expression study of Frattini *et al.* (43), which reveals similar global gene expression profiles of PTC harboring different genetic lesions. Although the molecular etiology of about 70% of all PTC is characterized by activating mutations or chromosomal rearrangements of *BRAF*, *RET*, or *RAS* (43, 61–63), a changed expression pattern of genes associated with the RAS-BRAF-MAPK pathway has not been described in human PTC microarray studies (31–33, 38, 40, 42–46). However, in a recent reanalysis of the data sets of Huang *et al.* (38) and Jarzab *et al.* (44) we did identify a significantly increased expression pattern for various genes of this pathway by GenMAPP analysis (48). Moreover, gene expression studies of papillary carcinoma cell lines harboring the *RET/PTC3* rearrangement, a mutant *HRAS* (v-Ha-ras) or a mutant *BRAF* oncogene attributed the differential expression of numerous PTC markers, like *LGALS3* and *DUSP6*, to the activation of the RAS-BRAF-MAPK pathway (58).

Although gene expression profiles of PTC with different genotypes (*i.e.*, *RAS* or *BRAF* mutation, *RET/PTC* rearrangements) share a high similarity, they are not identical (Table 2). Melillo *et al.* (58) could show that a large number of genes are specifically modulated by only one or two of the three oncogenes (*i.e.*, *RET/PTC*, *RAS*, and *BRAF*). Therefore, the three oncoproteins, which are part of a single signaling pathway, are each able to trigger specific signals in addition to the common ones. This conclusion is supported and widened by the gene expression profiling of Giordano *et al.* (46), who could show strong relationships not only between gene expression and PTC morphology, but also between the gene expression pattern and the occurrence of *RET/PTC* rearrangements and *BRAF* or *RAS* mutations. Therefore, it is very likely that the activation or interaction of/with additional/alternative signaling pathways forms the molecular basis of tumors and causes discrete mutation-specific phenotypic features (*e.g.*, morphology).

TABLE 2. Common and distinct genes for the PTC subtypes characterized by *RET/PTC*, *BRAF*, and *RAS* mutations

PTC subtype	Distinct genes
RET/PTC	Up-regulated: <i>RET</i> , <i>KIAA0894</i> , <i>MYL4</i> , <i>FLJ14152</i> , <i>WARP</i> , <i>Hs.116678</i> , <i>FLJ20949</i> , <i>C21orf106</i> , <i>FLJ20716</i> , <i>TNXB</i> , <i>DKFZp586A061</i> , <i>IGF1</i> , <i>FLJ10534</i> , <i>FLJ13829</i> , <i>LOC440423</i> , <i>DAPP1</i> , <i>CX3CL1</i> , <i>SLC35E1</i> , <i>SMA4</i> , <i>ZNF611</i> Down-regulated: <i>LOC387869</i> , <i>FLOT1</i> , <i>YWHAE</i> , <i>PAX8</i> , <i>FLJ10074</i> , <i>GGA2</i> , <i>HSPC121</i> , <i>C6orf106</i> , <i>MAX</i> , <i>ENSA</i> , <i>RAB5A</i> , <i>C19orf6</i> , <i>AFURS1</i> , <i>UBE4B</i> , <i>MBNL2</i> , <i>RNP24</i> , <i>HNRPU</i> , <i>DHX9</i> , <i>NUMA1</i> , <i>PRPF4B</i>
BRAF	Up-regulated: <i>TM7SF4</i> , <i>FN1</i> , <i>PLXNC1</i> , <i>MET</i> , <i>ARNTL</i> , <i>TMPRSS6</i> , <i>PERP</i> , <i>GATA3</i> , <i>IRS2</i> , <i>SLC34A2</i> , <i>PDE5A</i> , <i>TPD52L1</i> , <i>ERBB3</i> , <i>CLECSF2</i> , <i>STAT1</i> , <i>ODZ3</i> , <i>GPLY</i> , <i>MLPH</i> , <i>ITGB6</i> , <i>BHLHB2</i> Down-regulated: <i>HGD</i> , <i>GALNT10</i> , <i>CACNA202</i> , <i>CA12</i> , <i>FHL1</i> , <i>HES1</i> , <i>VAV3</i> , <i>ARGBP2</i> , <i>DUOX1</i> , <i>TFF3</i> , <i>PDE4A</i> , <i>IRS1</i> , <i>IGFBP3</i> , <i>MMP15</i> , <i>HLF</i> , <i>COL4A5</i> , <i>MAGED4</i> , <i>PLA2G4B</i> , <i>SLIT1</i> , <i>ARHN</i>
RAS	Up-regulated: <i>SLC4A4</i> , <i>HIST1H4H</i> , <i>CYP27B1</i> , <i>PIWIL1</i> , <i>SCGB2A1</i> , <i>KCNAB1</i> , <i>C16orf33</i> , <i>GNA14</i> , <i>DJ167A19.1</i> , <i>TRPC6</i> , <i>NQO1</i> , <i>DCI</i> , <i>C1orf41</i> , <i>DIO1</i> , <i>CTAGE5</i> , <i>GPR30</i> , <i>BCL2</i> , <i>ABCC8</i> , <i>HHEX</i> , <i>PTPN21</i> Down-regulated: <i>FN1</i> , <i>COL8A2</i> , <i>CFH</i> , <i>PAM</i> , <i>SLC34A2</i> , <i>LGALS1</i> , <i>KCNN4</i> , <i>EMP3</i> , <i>DUSP6</i> , <i>TIMP1</i> , <i>MET</i> , <i>SERPINA1</i> , <i>CFHL1</i> , <i>TRIO</i> , <i>HIC</i> , <i>MGC3047</i> , <i>PDLIM4</i> , <i>S100A10</i> , <i>CTSC</i> , <i>TACSTD2</i>
	Common genes
RET/PTC, BRAF, RAS	Up-regulated: <i>Adm</i> , <i>Dusp6</i> , <i>Itga1</i> , <i>Marks</i> , <i>Igf2r</i> , <i>Sgk</i> , <i>PI4K2B</i> , <i>Fhl2</i> , <i>CITED2</i> , <i>IRF7-like</i> , <i>Spp1</i> , <i>CCL2</i> , <i>CXCL1</i> , <i>Mmp13</i> , <i>Mmp10</i> , <i>Mmp12</i> , <i>Mmp3</i> , <i>USP18-like</i> , <i>GSTM2</i> , <i>Th</i> , <i>Fdps</i> , <i>Ca2</i> , <i>Ddbi</i> , <i>Ldlr</i> , <i>Lgals3</i> , <i>Col1a1</i> , <i>Tmsb4x</i> , <i>Vim</i> , <i>Lgals1</i> , <i>Dysf2-like</i> , <i>S100A4</i> , <i>S100A6</i> Down-regulated: <i>Gas6</i> , <i>Nr4a1</i> , <i>Crem</i> , <i>Hhex</i> , <i>Nr4a3</i> , <i>IRF8</i> , <i>Titf1</i> , <i>Tpo</i> , <i>Tshr</i> , <i>Tg</i> , <i>Dio1</i> , <i>Trg</i> , <i>Slc5a5</i> , <i>Knk3</i> , <i>Decr1</i> , <i>Ephx1</i> , <i>Pygb</i> , <i>Ass</i> , <i>Knk16</i> , <i>Rdh10</i> , <i>Knk1</i> , <i>Vdup1</i> , <i>GCG</i> , <i>ACY1</i> , <i>Spna2</i> , <i>Ibsp</i> , <i>Sdc2</i> , <i>Acta1</i> , <i>WASL</i> , <i>Gira1</i> , <i>Pde4b</i> , <i>Pde4d</i> , <i>Pgf</i> , <i>Adra1b</i> , <i>Pla2g4a</i> , <i>Itpr2</i> , <i>Lrp2</i> , <i>Ppp1r1b</i> , <i>Pak3</i> , <i>Rdc1</i> , <i>IRS-2</i> , <i>Mig6</i> , <i>PEL1</i> , <i>PI4P5K-1B</i> , <i>F10</i> , <i>TCRVAQ4a23</i> , <i>Knq</i> , <i>Dig-1</i> , <i>LOC207125</i> , <i>Tcp1</i>

Distinct genes are from Ref. 46, and common genes are from Ref. 58.

Although the correlation of *BRAF* mutations with a more aggressive form of thyroid carcinoma is discussed controversially (64–70), the observation of decreased *TPO* expression in *BRAF* mutant PTC (46) may also be of clinical relevance. Because *TPO* plays a central role in thyroid hormone synthesis, *BRAF* mutant PTC would have less radioiodine uptake and could thus be less responsible to radioiodine therapy (46). If this holds true, these observations would also be in line with the observation of a less favorable prognosis of *BRAF* mutant PTC (64, 69, 70).

Differences in the gene expression profile of PTC variants were indicated by Giordano *et al.* (46), who showed that PTCs of classic, follicular, and tall cell variants did cluster separately by principal component analysis. The difference between classic and tall cell variant was also described by Wreesmann *et al.* (47). On the other hand, the differences between the classic and follicular PTC variant are less visible (40), although some authors published lists of genes differentiating both variants (32).

The various groups specify different markers for the molecular diagnosis of PTC, such as *SFTPB* and *CITED 1* (38), *ADM* (adrenomedullin), *TROP2* (tumor-associated calcium signal transducer 2), and *NRP2* (neuropilin 2) (40); the gene set of *KIT* (v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog), *CDH1* (cadherin 1), *LSM7* (LSM7 homolog, U6 small nuclear RNA associated), *SYNGR2* (synaptogyrin 2), *FAM13A1* (family with sequence similarity 13, member A1), *IMPACT* (Impact homolog) and four nameless genes (42); and a 20-gene classifier (44) as well as gene lists comprising hundreds of genes (33, 40). This issue will be addressed below.

### B. Follicular thyroid carcinoma (FTC)

Gene expression profiles of FTC were analyzed in several microarray studies, with the aim to improve the molecular

differentiation between FTC and follicular adenomas and to elucidate the molecular etiology of FTC further (27–37).

In an early investigation, Barden *et al.* (29) compared the gene expression profiles of FTC and follicular adenoma using Affymetrix GeneChips and found strikingly distinct gene expression patterns for both entities. They identified 105 differentially expressed genes between FTC and follicular adenoma including *ADM*, *GPC1* (glypican 1), *TGFA* (TGF $\alpha$ ), *MET* and *IGFBP3* (IGF binding protein 3) (up-regulated in FTC), and *DIO1*, *CREM* (cAMP-responsive element modulator), *FBLN5* (fibulin 5), and *MT1* (metallothionein 1) (down-regulated in FTC). An alternative approach of microarray analysis that aims at identifying the minimal number of discriminating genes appears more promising for diagnostic purposes and could also lead to further elucidation of the molecular etiology. Weber *et al.* (30) applied the linear discriminant analysis to their data set of FTC and benign thyroid samples to find the best combination of genes discriminating the different entities. Using this approach, they identified the combination of *CCND2* (cyclin D2), *PCSK2* (proprotein convertase subtilisin/kexin type 2), and *PLAB* (growth differentiation factor 15) that allowed accurate differentiation between benign and malignant follicular thyroid neoplasia with a high sensitivity and specificity. Cerutti *et al.* (71) have specified another set of genes, based on serial analysis of gene expression (SAGE) analysis, which comprised *DDIT3* (DNA-damage-inducible transcript 3), *ARG2* (arginase, type II), *ITM1* (integral membrane protein 1), and *C1orf24*. Although these genes were primarily selected for diagnostic purposes (with the aim to differentiate between follicular adenomas and carcinomas), they should also be considered for their significance regarding the biology of FTC.

By a combination of microarray data and previously published loss of heterozygosity data of FTC and normal thyroid samples, Aldred *et al.* (27) identified five new putative tumor suppressor genes. Three of them belong to a set of most consistently down-regulated genes [*CAV1* (caveolin 1), *CAV2* (caveolin 2), and *GDF10* (growth differentiation factor 10)], whereas *GPC3* (glypican 3) and *CHRD1* (chordin-like 1) might be functionally related to *GDF10*. Although both caveolins are down-regulated in FTC, their molecular mechanisms of down-regulation remain unclear. Loss of heterozygosity was found in a subgroup of tumors, but the authors could not identify mutations in either gene, and furthermore, the methylation status of the caveolin-1 promoter did not correlate with the expression pattern (27). The investigation of *GDF10*, *GPC3*, and *CHRD1* revealed similar results. Moreover, *GPC3* and *CHRD1* have also been shown to be down-regulated in AFTNs (52) and CTNs (54). Therefore, these genes most likely do not make it possible to distinguish between benign and malignant thyroid tumors.

Subsequent to the independent analysis of their PTC and FTC microarray data, Huang *et al.* (38) and Aldred *et al.* (27) combined their data sets and compared the gene expression profiles of FTC and PTC to understand better the similarities and differences between these two tumor types (31). They identified genes that are characterized by a differential expression between FTC and PTC [*CST6* (cystatin E/M), *DPYSL3* (dihydropyrimidinase-like 3), *G0S2* (G0/G1switch 2), *IGFBP6* (IGF binding protein 6), *CITED1*, *CLDN10* (clau-

din 10), *CAV1*, and *CAV2*], and additional 28 genes including *TG* (thyroglobulin), *TPO*, *GPC3*, *DUSP1*, *APOD* (apolipoprotein D), *MT1G* (metallothionein 1G), and *CRABP1* (cellular retinoic acid binding protein 1) that are characterized by a decreased expression in both entities, which might suggest a common origin of FTC and PTC. However, this conclusion might be overstated because *TG*, *TPO*, *MT1G*, and *CRABP1* are characterized by an increased expression in AFTNs (52), which rather suggests a functional relevance for thyroid hormone production. In contrast, the decreased expression of *GPC3* and *APOD*, which has also been shown in AFTNs (52) and CTNs (54), most likely reflects increased proliferation that is inherent to all tumor entities.

In addition to activating *RAS* mutations, which have been found in follicular neoplasms (72), a fusion oncogene between *PAX8* and *PPAR $\gamma$*  has been identified in up to 63% of FTC (73), but also in a number of follicular adenomas (74–76). However, the knowledge about the detailed molecular events by which *PAX8-PPAR $\gamma$*  contributes to tumor development is so far limited (77). Therefore, to identify molecular signatures of FTC with and without the *PAX8-PPAR $\gamma$*  fusion oncogene, Lui *et al.* (36) performed a gene expression profiling of both subtypes. In their study, Lui *et al.* could show a highly uniform expression profile in the FTC characterized by a *PAX8-PPAR $\gamma$*  fusion oncogene that differed significantly from those of the FTC without the fusion oncogene. On the basis of these results, they reasoned that FTCs with a *PAX8-PPAR $\gamma$*  fusion oncogene form a distinct biological entity within the FTC. Moreover, this conclusion is in line with other findings. French *et al.* (78) demonstrated distinct clinicopathological features in FTC with a *PAX8-PPAR $\gamma$*  rearrangement. Nikiforova *et al.* (79) proposed two distinct pathways involved in the development of FTC (initiated by either *RAS* point mutations or *PAX8-PPAR $\gamma$*  rearrangements). A gene expression profiling study of Lacroix *et al.* (34) also identified specific gene expression patterns associated with the *PAX8-PPAR $\gamma$*  rearrangement. In addition, the assumption of a distinct *PAX8-PPAR $\gamma$*  pathway is supported by a differential expression of signal transduction genes [*e.g.*, increased: *EPOR* (erythropoietin receptor), *ADCY9* (adenylate cyclase 9), *EPAC* (Rap guanine nucleotide exchange factor 3), and *DUSP6*; decreased: *HINT1* (histidine triad nucleotide binding protein 1), *RAP1GAI* (RAP1 GTPase activating protein), and *CSNK2B* (casein kinase 2)] and proliferation-associated genes [*e.g.*, *CTBP2* (C-terminal binding protein 2), *FSCN1* (fascin homolog 1), and *CCND1* (cyclin D1)] in FTC with a *PAX8-PPAR $\gamma$*  rearrangement (36). Recently, Giordano *et al.* (28) studied the implications of the *PAX8-PPAR $\gamma$*  fusion protein for the neoplastic mechanism of FTC by a comprehensive investigation of gene expression profiles of *PAX8-PPAR $\gamma$* -positive FTC in comparison to *PAX8-PPAR $\gamma$* -negative FTC, along with other common thyroid tumors such as follicular adenomas, oncocytic carcinomas, oncocytic adenomas, and papillary carcinomas. *PAX8-PPAR $\gamma$* -positive FTC could be identified by examining the microarray data for increased expression of *PPAR $\gamma$*  (28). Moreover, principal component analysis of all follicular neoplasms revealed a separation of the *PAX8-PPAR $\gamma$* -positive FTC from the *PAX8-PPAR $\gamma$* -negative FTC and the follicular adenomas, suggesting that the *PAX8-PPAR $\gamma$*  rearrangement is the predominant

source of variation in gene expression within this set of tumors (28). Interestingly, in their study Giordano *et al.* (28) could show that the PAX8-PPAR $\gamma$  fusion protein can function in a PPAR $\gamma$ -like manner, although it also has transcriptional properties distinct from either PAX8 or PPAR $\gamma$ . These findings are contrary to the original description of the PAX8-PPAR $\gamma$  fusion protein (73). Kroll *et al.* (73) showed that the fusion protein does not activate PPAR-responsive promoters, but functions as a dominant-negative inhibitor of PPAR $\gamma$ -induced reporter gene activation. Therefore, the inhibition of endogenous PPAR $\gamma$  is said to be an important mechanism by which the PAX8-PPAR $\gamma$  fusion protein causes FTC. However, due to Giordano's findings (28), the concept that the PAX8-PPAR $\gamma$  fusion protein contributes to follicular carcinoma by antagonizing endogenous PPAR $\gamma$  needs re-evaluation. Advanced bioinformatic analysis of Giordano's data revealed potential roles of several metabolic pathways in the oncogenic action of the fusion protein. An enrichment of the PAX8-PPAR $\gamma$  expression signature was observed in pathways related to fatty acid oxidation and metabolism, and to amino acid and carbohydrate metabolism. These results are striking because PPAR $\gamma$  regulates adipogenesis and glucose metabolism.

### C. Autonomously functioning thyroid nodules (AFTNs)

In an early study using membrane arrays and radioactively labeled cDNA probes, we could show a down-regulation of several signal-transducing components such as *IGF-II*, the platelet-derived growth factor receptor (*PDGFR*) and the type III TGF- $\beta$  receptor (*TGFBR3*) in AFTNs and CTNs, which suggested a disturbed signaling system (8). In a recent study, we confirmed these findings. The TGF- $\beta$  signaling cascade is particularly characterized by strong changes in its pattern of gene expression in AFTNs when compared with their normal surrounding tissues (ST): the *TGFBR3*, *SMAD 1*, *3*, and *4*, as well as p300, a transcriptional coactivator, showed a decreased expression in AFTNs, whereas the inhibitory *SMAD 6* and *7* showed an increased expression in AFTNs. Moreover, a decreased expression of *TGFBR3* and *TGFB1* could also be shown at the protein level (52, 80). These findings suggest an inactivation of TGF- $\beta$  signaling in AFTNs due to a constitutively activated TSH receptor (TSHR) (*e.g.*, resulting from *TSHR* mutations) (52).

Because the molecular etiology of AFTNs is mainly characterized by a constitutively activated TSHR-cAMP-pathway, we compared the gene expression profiles of TSH-stimulated primary thyroid epithelial cells (51). Because these data also showed significant differences in the gene expression profiles of the TGF- $\beta$  signaling cascade, the findings in the AFTNs are most likely due to the constitutively activated cAMP cascade. This is further supported by investigations of TSH-stimulated porcine thyroid follicles *ex vivo* (81) showing a decreased *TGFB1* mRNA expression. Furthermore, in the context of an increased  $\beta$ -arrestin 2 expression in AFTNs (82), findings of a  $\beta$ -arrestin 2 mediated endocytosis of the *TGFBR3* and a subsequent down-regulation of its signaling by Chen *et al.* (83) support and further elucidate this interpretation. It has been suggested that a down-regulation of lymphocyte and macrophage-specific genes

(52, 53) might reflect a different cellular composition of the AFTNs with a lack of lymphocytes and discussed by Wattel *et al.* (53). However, the specific expression patterns of TGF- $\beta$  signaling-associated genes in thyrocytes are not influenced by the different cellular composition and could be reproduced in TSH-stimulated primary thyroid epithelial cells (51).

Interestingly, two independent studies using different microarray platforms (*i.e.*, spotted glass slides and Affymetrix GeneChips) and different methods of data analysis consistently show an increased expression of thyroid-specific genes such as *TPO* and *DIO1*, as well as collagen (type IX,  $\alpha 3$ ), different metallothioneins, and sialyltransferase (*SIAT1*) in AFTNs (52, 53). The stringent and prominent *SIAT1* expression pattern that we found in our microarray study of AFTNs (52) prompted us to investigate one of its possible functional relevances further. Subsequent studies characterized a new aspect of posttranslational modification of the TSHR. In cell culture experiments, we demonstrated for the first time that the transfer of sialic acid directly affects TSHR signaling because it improves and prolongs the cell-surface expression of the TSHR (84). Furthermore, microarray investigations of AFTNs and TSH-stimulated primary thyroid cells illustrate a remarkable induction or activation of negative feedback mechanisms such as an up-regulation of phosphodiesterases (11, 51, 53), which is in line with previous findings of Persani *et al.* (85). Moreover, the presence of negative feedback mechanisms in AFTNs is further supported by an increased expression of G protein-coupled receptor kinases (GRK) in AFTNs and their ability to desensitize the TSHR as shown in *in vitro* experiments (86). Interestingly, whereas microarray studies of TSH-stimulated primary thyroid cell cultures reveal an increased expression of the regulator of G protein signaling (*RGS 2*) (11, 51), which has been shown to reduce the TSHR signaling via inositol-3-phosphate (87), *RGS2* is characterized by a decreased expression in AFTNs (11, 51, 88). Such a difference can most likely be explained by defects in the *RGS* regulation pathway or by additional counter-regulatory mechanisms that occur only in the chronically proliferating AFTNs. Furthermore, in TSH-stimulated primary thyroid cell cultures, an increased expression of *CREM*, which acts as a repressor of cAMP-induced genes (89), could be observed, although this was not observed by microarray studies of AFTNs (11, 51). These findings demonstrate that the additional investigation of a compatible cell model (*e.g.*, TSH-stimulated primary thyroid epithelial cells) may confirm and further define and/or explain specific gene expression patterns (*e.g.*, TGF- $\beta$  signaling cascade) detected in tissue samples (*e.g.*, AFTNs). However, it is obvious that the gene expression profile of the cell model itself can only depict a part of the complex situation present in the tissue samples.

The comparison of the gene expression patterns of AFTNs harboring a somatic TSHR mutation and TSHR mutation-negative AFTNs (52) indicated a number of signal transduction genes [*e.g.*, p21/Cdc42/Rac1-activated kinase (*PAK*)1 and 2, *RGS4* and -6, Janus kinase (*JAK*)1, and G protein-coupled receptor kinase (*GRK*)2] which are characterized by a different expression pattern between these two subgroups of AFTNs. These findings indicate strong differences in the

signal transduction of AFTNs caused by constitutively activating mutations in the TSHR compared with AFTNs without TSHR mutations and could thus give a lead for the elucidation of their molecular etiology.

Taken together, the global analysis of gene expression profiles of AFTNs showed both an inactivation of TGF- $\beta$  signaling and a remarkable induction or activation of negative feedback mechanisms in the AFTNs, which could also be confirmed in a compatible cell model in TSH-stimulated primary thyroid epithelial cells. Moreover, the prominent *SIAT1* expression pattern in AFTNs provided the basis for the characterization of a new aspect of posttranslational modification of the *TSHR*.

#### D. Cold thyroid nodules (CTNs)

Several microarray studies, which were performed with the aim to identify novel diagnostic and clinical markers for differentiated thyroid tumors, used benign thyroid nodules, histologically classified as follicular adenoma or hyperplastic nodules, for comparison with PTC and FTC (29, 30, 32, 33, 35, 37, 40, 42, 49, 90). However, with respect to their function, histologically benign thyroid nodules can be further distinguished as AFTNs or CTNs or less often as so-called warm nodules, which do not show detectable differences to the surrounding thyroid tissue by scintigraphy. CTNs constitute the most abundant thyroid nodular lesion with a prevalence of about 80% of all palpable thyroid nodules (91). In contrast to AFTNs (described in Section II.C), the current knowledge concerning their molecular etiology is very limited. Nevertheless, there is only one study that investigated the array-based gene expression profiles of CTNs classified by scintigraphy (54).

In an early study we could show a down-regulation of several signal transducing components both in AFTNs and CTNs, which seemed to reflect a disturbed signaling system (8). These similarities in the gene expression patterns of AFTNs and CTNs might be attributable to a common property of both benign tumor entities, *e.g.*, increased proliferation (92, 93). To gain a higher resolution that might help to identify specific signaling cascades involved in nodular development, we subsequently compared gene expression profiles of 22 CTNs to their normal surrounding tissue using the U95A Affymetrix GeneChip (54). On the basis of the high number of investigated genes (approximately 10,000 full-length genes) and an improved statistical analysis that made it possible to analyze the significance of differential gene expression within gene sets (*e.g.*, signaling cascades), the molecular pattern of the increased proliferation in CTNs (93) could be further defined. Regulation of gene expression was most consistent for a number of histone mRNAs and gene sets containing cell cycle-associated genes, like cyclin D1, cyclin H/cyclin-dependent kinase (*CDK*) 7, and cyclin B. Furthermore, these expression data also revealed that contrary to PTCs, altered expression of components belonging to the RAS-MAPK cascade is of minor importance for the development of CTNs because gene sets representing this pathway did not show differential expression in comparison to the surrounding normal tissue. This is in line with findings of Esapa *et al.* (94), who showed that a general relevance of

RAS mutations for the development of follicular adenoma is unlikely. Moreover, these results are supported by findings of Krohn *et al.* (95) and a recent *in vitro* study indicating that the dedifferentiated phenotype of CTNs is unlikely to be the result of an activated RAS signaling (96).

Interestingly, the gene set analysis also revealed a significantly altered expression pattern in the group of G protein signaling molecules, which is mainly based on a differential expression of several protein kinase C (*PKC*) isoforms and an increased expression of the  $G_q\alpha$  protein. This expression pattern is especially interesting because it has been shown that thyroid cells undergoing a long-term *PKC* stimulation are characterized by a general loss of thyroid-specific functions (*e.g.*, loss of iodide transport and thyroglobulin iodination), which is reminiscent of CTNs (97–101).

The up-regulation of cell cycle genes in the gene expression profile of CTNs (54) is so distinct that it may be used for the differential molecular diagnosis of these tumors, as shown by RFR algorithm (44). Mainly genes related to proliferation and growth processes were included into a 20-gene molecular classifier (*e.g.*, cyclin-dependent kinase inhibitor 1C, histone 1 H2be, and histone 2 H2aa). Furthermore, the CTN classifier contains genes considered as cancer specific, such as the fibroblast growth factor receptor 1 (*FGFR1*) found down-regulated by Chevillard *et al.* (32) in the follicular variant of PTC in comparison with the classic variant, *TLE4* (Transducin-like enhancer protein 4), described by Aldred *et al.* (27) as down-regulated in FTC, and *TUSC3* (tumor suppressor candidate 3) overexpressed in PTC (38). In contrast to AFTNs, the CTN multigene signature was also found in some PTC, which could be explained by the partly dedifferentiated and proliferating phenotype that is common to both entities.

Several studies provide evidence that differentiated functions of thyrocytes and of iodide metabolism can be reinduced by retinoic acid (102–104). Therefore, the significantly changed expression pattern of *CRABP1* in CTNs is of special interest (54). *CRABP1* encodes a high-affinity cellular retinoic acid binding protein that regulates the availability of retinoic acid for its nuclear receptors and is also involved in retinoic acid catabolism (105). Therefore, the decreased mRNA expression and the decreased protein expression of *CRABP1* in CTNs might also impact on the partly dedifferentiated and hypofunctioning phenotype of this tumor entity. Our recent findings showing an increased expression of *CRABP1* in AFTNs in comparison to their normal surrounding tissue support this assumption (unpublished observation). Moreover, because a decreased *CRABP1* expression has also been shown in PTC (38) and FTC (27) in addition to CTNs, a reduced expression of *CRABP1* most likely reflects an early dedifferentiating event in the pathogenesis of thyroid tumors.

Overall, in addition to the interesting finding of decreased *CRABP1* expression, gene expression profiling of CTNs allowed identification of the molecular pattern of their increased proliferation that is characterized by a differential expression of several cell cycle-associated genes. Moreover, whereas the RAS-MAPK cascade is most likely of minor importance for the development of CTNs, the gene expression patterns of the  $G_q$ -*PKC* pathway suggest further inves-

tigations to define its relevance in relation to the dedifferentiated phenotype of CTNs.

#### *E. Clues for the comparison and differential diagnosis of thyroid nodules, based on the microarray data*

Despite the fact that microarray studies revealed very distinct changes in the expression of certain genes, none of the several genes identified by array studies as differentially regulated was proven to be an ideal single marker of PTC (106–115). For example, *DPP4* (dipeptidyl-peptidase 4), which was indicated by Huang *et al.* (38) and later shown by Jarzab *et al.* (44) to be the most up-regulated gene in PTC, did not clearly differentiate between PTC and benign tissue (116). Also, oncofibronectin, galectin 3, and other proposed markers did not work properly in a single gene context (117–121). Moreover, the different multigene classifiers proposed by various authors showed only a minor overlap of the respective markers included. In fact, such a comparison requires a systematic bioinformatic approach. Such analysis has been performed in some types of cancer (122) but is not available for thyroid cancer.

In a study comprising 62 samples of PTC, follicular variant of PTC, FTC, medullary thyroid carcinoma, and follicular adenomas/hyperplastic nodules, a set of 662 differentially expressed genes was identified that discriminates between benign and malignant thyroid tumors (33). The discriminating gene set contained both known cancer-associated genes (*e.g.*, *LGALS3*, *TIMP*, *TGFA*, *ADM*, *MET*, and *FN1*) and previously unidentified genes. However, such large gene lists are not applicable for diagnostic purposes. Thus, approaches to limit the number of genes in an identifier have been undertaken. The aim of the study of Jarzab *et al.* (44) was not

to list genes with the largest fold-changes, but rather to specify the most powerful set of genes that discriminate PTC and benign tissue. This gene set was selected not by univariate approaches but by complementation of the genes, by a RFR algorithm. Interestingly, the classifier does not contain many known genes found in other approaches like *FN1* or *TIMP*, whereas some other genes previously known for their up-regulation in PTC (*e.g.*, *DPP4*, *SERPINA1*, *LGALS3*, and *MET*) were included (Fig. 1). Within the classifier there were some new genes, previously not described in PTC [*e.g.*, *EVA1* (epithelial V-like antigen 1), and *LPR4* (transmembrane 6 superfamily member 2)] or not evaluated in a diagnostic context (retinoid X receptor,  $\gamma$ , *RXRG*), as well as genes exhibiting less distinct changes in expression (*e.g.*, gap junction protein,  $\beta 3$ , *GJB3*). The goal of this approach was not to obtain 20 genes showing a correlated change in expression between tumor and benign tissue but to generate complementary information from a selected set of genes. The idea of using gene interactions to improve classification accuracy can be explained with the simplest example of a two-gene linear interaction (Fig. 2). Within the shown data set of papillary thyroid cancers and normal/benign thyroid tissues, there are samples that cannot be classified by single markers, but their class could be predicted by a combination of them. However, two genes cannot classify all samples, and it is necessary to base the classification on a larger number of genes, usually five to 20.

Taking into account the known heterogeneity of malignant tumors, RFR algorithm helped to obtain a robust molecular classifier able to recognize a wide range of PTC tumors. The efficacy of the obtained multigene classifier was evaluated

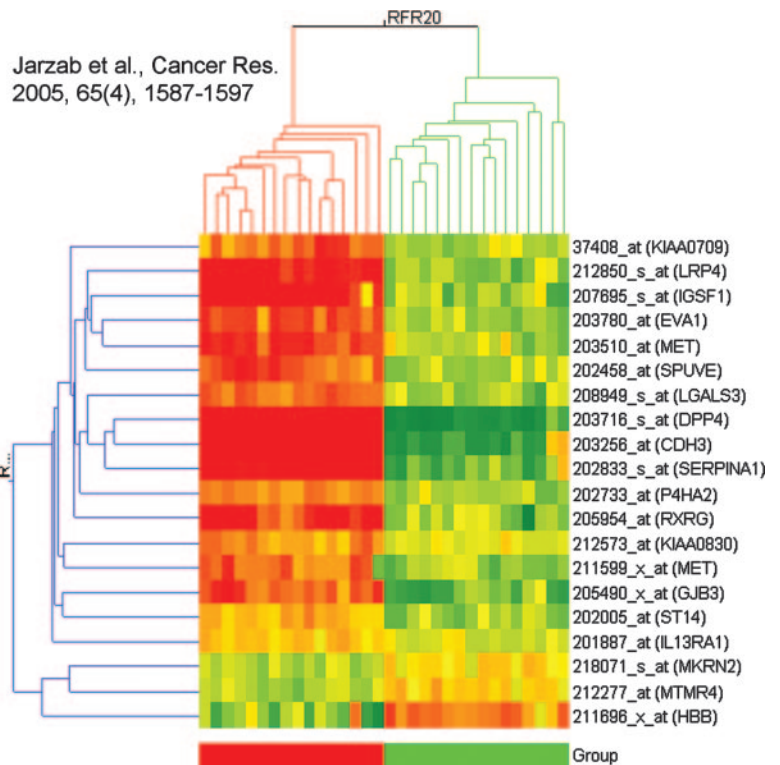


FIG. 1. The 20-gene PTC classifier obtained by RFR [the name RFR-20 refers to the number of genes included in the classifier (44)] performs very well to differentiate between tumors (*red bar*) and normal tissues (*green bar*). The intensity of the color refers to the intensity of gene expression; *red* refers to the up-regulation, and *green* refers to the down-regulation. However, only some genes (*e.g.*, dipeptidylpeptidase 4, *DPP4*, to a lesser extent cadherin P, *CDH3*, and  $\alpha 1$ -antitrypsin, *SERPINA 1*) show a stable and very distinct difference in expression between both types of tissues. Other included genes (*e.g.*, *MET* proto-oncogene, represented here by two different probe sets) give slightly different results. Also *RXRG* or galectin 3, *LGALS3*, are over-expressed only in a subset of PTC and, in the case of lack of overexpression, the information is completed by the expression of other genes.



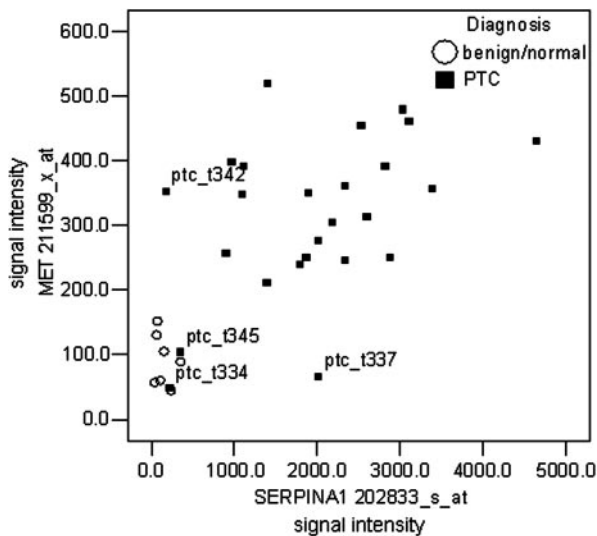


FIG. 2. Classification of thyroid samples into two classes (papillary thyroid cancer and normal/benign thyroid), based on gene expression data for two transcripts, *SERPINA1* and *MET*. Class prediction for sample ptc\_t342 is correct only when based on expression of *MET*, whereas for ptc\_t337 only *SERPINA1* classifies this sample correctly. Taken together, these two markers make it possible to classify both samples properly. For the other two samples (ptc\_t334 and 345), separation from normal samples based on these two genes is impossible, and further genes have to be included.

and confirmed in further studies, showing 93–95% accuracy (48).

A more formal analysis of the accuracy of molecular diagnosis of PTC was also conducted with SVMs (K. Fujarewicz, M. Jarzab, M. Eszlinger, K. Krohn, R. Paschke, M. Oczko-Wojciechowska, M. Wiench, A. Kukulka, B. Jarzab, and A. Swierniak, manuscript in preparation). It proved that 98% of accuracy may be achieved with 95% confidence and limits the range between 95 and 100%. The lower limit of the 95% confidence interval reached 95% already with classifiers composed of only five genes. In this context, it is advisable to proceed with molecular classifiers composed of at least six genes. It is more than a simple coincidence that Mazzanti *et*

*al.* (42) succeeded to differentiate their PTC samples from benign samples by a gene set of six genes and that another molecular classifier proposed for PTC on the basis of microarray data and published recently also consists of six genes (123).

Tumors and tissues often consist of more than one type of cells that might contribute differently to the measured expression of a given gene (124). The resulting problem of data interpretation in such a complex situation can be solved in different ways. Before the microarray experiment is performed, specific cell types can be isolated by microdissection. Alternatively, after microarray hybridization the expression of specific genes can be localized by *in situ* hybridization. Venet *et al.* (124) suggest a mathematic algorithm to separate samples consisting of different cell types. We prefer to include at least 10–20 genes into the classifier, both tumor markers and stromal cell markers. Still, at this number of genes, the conversion to a multi-quantitative PCR approach is feasible (125). Others prefer a wider minimalization, mainly because they intend to perform immunohistochemical identification. Weber *et al.* (30) have proposed a three-gene classifier for differentiation of follicular tumors, and Cerutti *et al.* (71) included four genes for the same purpose. These proposed marker sets should be reinvestigated in fine-needle aspiration (FNA) biopsies (FNABs) for their applicability to FNAB diagnosis.

### III. First Results, Possibilities, and Solutions for Data Set Integration and Meta-Analysis

#### A. Study design

The diverse reports investigating the gene expression profiles of both benign and malignant thyroid tumors show a highly variable study design. In addition to the use of different microarray platforms, such as Affymetrix GeneChips and spotted oligo or cDNA arrays, there are important methodological differences between these reports with respect to: 1) GeneChip generation [*e.g.*, U95A(v2) and U133A]; 2) reference tissue; and 3) the kind of array comparisons (Fig. 3).

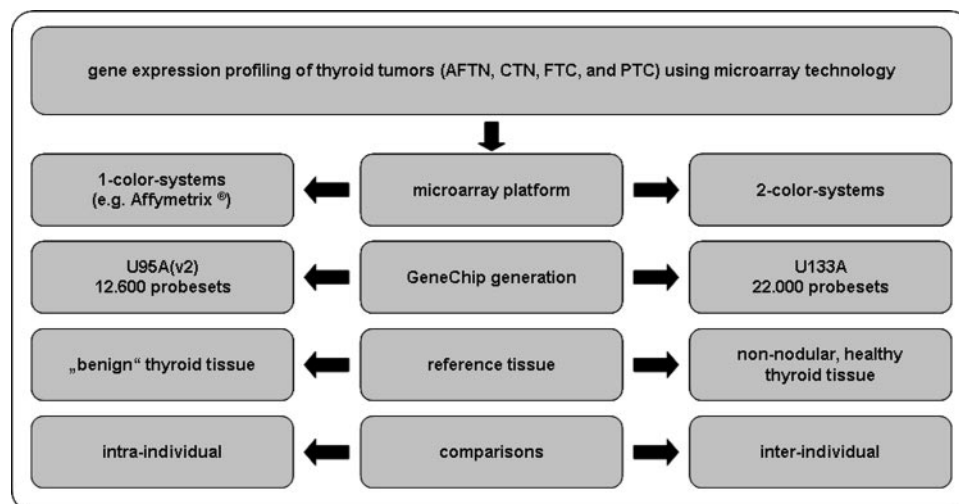


FIG. 3. Methodological heterogeneity of microarray studies investigating the gene expression profiles of thyroid tumors.

In addition, the diversity of the different studies is further increased by a large variety of applied algorithms for data preprocessing and data analysis. Although early studies and studies employing radioactively labeled macroarrays used housekeeping genes for data normalization (8) or simple algorithms such as z-score transformation (9, 10), the introduction of high-density microarrays such as the Affymetrix GeneChips goes along with the development of a number of much more sophisticated preprocessing algorithms (*e.g.*, MAS5, RMA, GC-RMA) that integrate information from up to 18 pairs of oligonucleotides (collectively known as a probe set) into a single expression value per RNA transcript. Furthermore, the analysis of the preprocessed expression values allows the application of a variety of mathematical procedures. These start with simple filtering to identify differentially expressed genes at the low end and comprise statistical procedures adjusting for multiple testing (*e.g.*, FDR, permutation procedures according to Westfall-Young, or significance analysis of microarrays), cluster analysis, and supervised learning techniques (*e.g.*, SVMs with RFR) at the high end. It is therefore necessary to compare the results of microarray studies in the context of the complexity of their generation, which does not mean that the more complex analysis ensures more meaningful results.

#### *B. Types and generations of microarrays (one-color systems vs. two-color systems) and GeneChip generations*

Both microarray platforms, *i.e.*, spotted cDNA/oligonucleotide microarrays and Affymetrix GeneChips, have been employed to an approximately equal extent in the analysis of gene expression profiles of both malignant thyroid tumors [FTC (27, 29–37) and PTC (31–33, 37, 38, 40–50)] and benign thyroid tumors [AFTNs (8, 11, 48, 51–53) and CTNs (8, 48, 51, 54)]. Due to the strong methodological and technological differences between these two platforms, concerns about the reproducibility and comparability of experimental results obtained with different platforms appear reasonable. However, whereas early studies found very poor correlations between different microarray platforms (126–128), recent studies show much more promising results and indicate that microarray data can actually be reproducible and comparable between different platforms and laboratories (129–131). Nevertheless, because of the rapid progress in microarray technology, a problem emerges that concerns comparability between microarray platforms. In a recent study, a comparison of gene expression profiles of AFTNs, CTNs, and PTC was performed; the samples were hybridized to two generations of Affymetrix GeneChips (AFTNs/CTNs, U95Av2; PTC, U95A/U133A), and the expression data were joined based on Affymetrix Best Match Comparison Spreadsheet (for details see [www.affymetrix.com/support/technical/comparison\\_spreadsheets.affx](http://www.affymetrix.com/support/technical/comparison_spreadsheets.affx)). This method of preprocessing retained the differences between the microarray platforms; the difference between the expression patterns of PTC samples hybridized to different Affymetrix GeneChip generations (as assessed by principal component analysis) was larger than between the expression patterns of the different tissue entities (48). This problem was solved by building classifiers from each dataset separately and comparing them.

More sophisticated approaches are being developed to construct molecular classifiers from data obtained with different GeneChip generations. A recent study (K. Fajarewicz, M. Jarzab, M. Eszlinger, K. Krohn, R. Paschke, M. Oczko-Wojciechowska, M. Wiench, A. Kukulska, B. Jarzab, A. Swierniak, manuscript in preparation) analyzed a 180-array dataset [HG-U95A(v2) and HG-U133A GeneChips] that included a collection of benign thyroid tumors, normal tissues, and PTC by a SVM-based selection that was extended by bootstrapping. In this way, a reliable estimate of classification accuracy of benign/malignant classification was obtained with appropriate confidence intervals. Using the bootstrap-based approach, it was possible to calculate a multigene classifier with a prediction efficiency higher than 95% from the data derived from two different array generations and to rank genes according to their classification ability (Table 3).

In contrast to a direct comparison and metaanalysis of the expression data, Griffith *et al.* (132) recently presented an interesting “meta-review” method for ranking genes based on their published evidence. Based on the assumption that biologically relevant genes will be overrepresented and system-specific spurious genes underrepresented in different studies, their approach involved a “vote-counting” strategy based on the number of studies reporting a gene as differentially expressed and further ranking based on total sample size and average fold-change. A comparison of their meta-review method (using published gene lists) to a meta-analysis of a smaller subset of studies (for which raw data were available) showed a strong level of concordance. Therefore, because raw data are currently unavailable for about 60% of the studies, their approach represents an interesting alternative to identify consistent gene expression markers across different studies.

#### *C. Selection of reference tissue and intra- vs. interindividual comparisons*

Some microarray studies of thyroid tumors tried to answer (patho-) physiological questions or aimed to elucidate the molecular etiology of the tumor, whereas others looked primarily for genetic markers that could improve the differential diagnosis of thyroid tumors. Because of these different questions to be answered by the microarray experiments, the selection of the reference tissues and the way of comparing the expression data varied significantly. In some studies, intraindividual comparisons were performed, whereas other studies performed interindividual comparisons. Furthermore, in some studies the reference tissue was defined as strictly nonnodular healthy tissue (8, 32, 34, 37, 38, 44–46, 48, 51, 52, 54), whereas it also comprised benign lesions such as goiter, follicular adenoma, and hyperplastic nodules in other studies (29, 30, 32, 33, 40, 42, 90). Moreover, there are differences in the nomenclature of benign tumors: whereas we prefer a nomenclature that is based on the function of the benign tumors (AFTNs, CTNs), other reports base their nomenclature on the histological morphology of the tumors (follicular adenoma, hyperplastic nodules). Therefore, in the case of comparing different microarray studies or performing a meta-analysis, it is mandatory to keep these charac-

TABLE 3. Ranking of PTC genes, differentiating between PTC and benign thyroid tissues by the BBFR approach<sup>a</sup>

Rank	Gene symbol	Gene name	OMIM	Ref.
1	<i>MPPED2</i>	Metallophosphoesterase domain containing 2	600911	27, 31, 42
2	<i>HBA1/HBA2</i>	Hemoglobin, $\alpha$ 1/ hemoglobin, $\alpha$ 2	141800/ 141850	
3	<i>MET</i>	Met proto-oncogene (hepatocyte growth factor receptor)	164860	29, 33, 40, 44, 46, 155–157
4	<i>FN1</i>	Fibronectin 1	135600	29, 33, 40, 44, 46, 155, 156, 158, 159
5	<i>GALE</i>	UDP-galactose-4-epimerase	606953	
6	<i>QPCT</i>	Glutaminyl-peptide cyclotransferase (glutaminyl cyclase)	607065	29, 32, 33, 40, 44
7	<i>NELL2</i>	NEL-like 2 (chicken)	602320	29, 33, 40
8	<i>PGCP</i>	Plasma glutamate carboxypeptidase		27, 29–31, 33, 40
9	<i>DPP4</i>	Dipeptidylpeptidase 4 (CD26, adenosine deaminase complexing protein 2)	102720	38, 44, 156, 160, 161
10	<i>ADORA1</i>	Adenosine A1 receptor	102775	27, 31, 156
11	<i>HMG2</i>	High-mobility group AT-hook 2	600698	41, 50, 162
12	<i>RYR1</i>	Ryanodine receptor 1 (skeletal)	180901	29, 33, 40
13	<i>CDH16</i>	Cadherin 16, KSP-cadherin	603118	
14	<i>GJB3</i>	Gap junction protein, $\beta$ 3, 31 kDa (connexin 31)	603324	44
15	<i>EMID1</i>	EMI domain containing 1	608926	29, 33, 40, 71
16	<i>NRIP1</i>	Nuclear receptor interacting protein 1	602490	29, 33, 40
17	<i>MET</i>	Met proto-oncogene (hepatocyte growth factor receptor)	164860	29, 33, 40, 44, 46, 155–157
18	<i>DTX4</i>	Deltex 4 homolog ( <i>Drosophila</i> )		44, 156
19	<i>RAB27A</i>	RAB27A, member RAS oncogene family	603868	29, 30, 33, 40
20		CDNA clone IMAGE:4152983		
21	<i>BCL2</i>	B cell CLL/lymphoma 2	151430	41, 46, 47, 50, 156, 162
22	<i>TACSTD2</i>	Tumor-associated calcium signal transducer 2	137290	46
23	<i>DIO1</i>	Deiodinase, iodothyronine, type I	147892	8, 29, 33, 38, 40, 46, 47, 52, 156
24	<i>ITPR1</i>	Inositol 1,4,5-triphosphate receptor, type 1	147265	29, 33, 40, 47, 156, 159
25	<i>HBB</i>	Hemoglobin, $\beta$ /// hemoglobin, $\beta$	141900	44, 71, 161
26	<i>SNED1</i>	Sushi, nidogen and EGF-like domains 1		
27	<i>AHR</i>	Aryl hydrocarbon receptor	600253	29, 33, 40, 155
28	<i>HGD</i>	Homogentisate 1,2-dioxygenase (homogentisate oxidase)	607474	27, 29, 29, 31, 33, 38, 40, 46, 52, 156
29	<i>RXRG</i>	Retinoid X receptor, $\gamma$	180247	44, 163
30	<i>CA4</i>	Carbonic anhydrase IV	114760	29, 30, 33, 40
31	<i>SDC4</i>	Syndecan 4 (amphiglycan, ryudocan)	600017	29, 32, 33, 40, 156
32	<i>ENTPD1</i>	Ectonucleoside triphosphate diphosphohydrolase 1	601752	30
33	<i>TPO</i>	Thyroid peroxidase	606765	29, 33, 40, 52, 71
34	<i>KRT19</i>	Keratin 19	148020	29, 32, 33, 40, 44, 156
35	<i>ID3</i>	Inhibitor of DNA binding 3, dominant negative helix-loop-helix protein	600277	
36	<i>RUNX1</i>	Runt-related transcription factor 1 (acute myeloid leukemia 1; aml1 oncogene)	151385	
37	<i>LMOD1</i>	Leiomodin 1 (smooth muscle)	602715	
38	<i>RAB27A</i>	RAB27A, member RAS oncogene family	603868	29, 30, 33, 40
39	<i>FBXO9</i>	F-box protein 9	609091	
40	<i>TRIM58</i>	Tripartite motif-containing 58		
41	<i>MT1G</i>	Metallothionein 1G	156353	41, 50, 52, 156, 162
42	<i>ICAM1</i>	Intercellular adhesion molecule 1 (CD54)	147840	164

<sup>a</sup> K. Fujarewicz, M. Jarzab, M. Eszlinger, K. Krohn, R. Paschke, M. Oczko-Wojciechowska, M. Wiench, A. Kukulska, B. Jarzab, A. Swierniak, manuscript in preparation.

teristics of the different benign tumors in mind. Otherwise, discrepant and misleading results are preprogrammed.

Not only do advanced analyses such as comparison and meta-analysis of microarray studies require knowledge concerning the selection of the reference tissues and the type of comparison (*i.e.*, intra- or interindividual comparisons). In a recent study, we addressed the relevance of paired (intra-individual) *vs.* unpaired (interindividual) data sets (48). The comparison of paired and unpaired PTC data sets clearly indicated a higher quality of paired data sets. Despite the fact that the paired data sets were hybridized to different GeneChip generations, the paired data sets showed more similarities in their gene expression profiles than the paired and the unpaired data sets that were hybridized to the same

GeneChip generation and were generated in the same laboratory under identical conditions (however, there were some differences in the number of samples, which could also be relevant). The GenMAPP analysis of the paired PTC data sets showed significant differences in the expression pattern of the MAPK cascade, whereas the unpaired data set provided no indications for this alteration (48). Taken together, because intraindividual comparisons share more similarities even across GeneChip generations than intra- and interindividual comparisons using the same microarray generation, the elimination of the “individual background” by intraindividual comparisons seems mandatory, especially in studies whose objective is the identification of subtle differences, *e.g.*, to elucidate the etiology of a specific pathology. In con-

trast, in studies looking for diagnostic markers of malignant tumors such as follicular and PTC, the variability of reference tissues might be of less importance; on the contrary, it may help to find out more robust cancer markers.

Furthermore, gene expression of thyroid tissue that is considered normal with regard to tumor characteristics could vary considerably. Bruno *et al.* (133) demonstrated differences in the expression of the sodium iodide symporter (*SLC5A5*) and the iodide transporter (*SLC5A8*) between thyroid tissue stimulated with low and normal serum TSH levels. Because the identification of differentially regulated genes critically depends on the type or metabolic status of the reference tissue as described above, we recently focused on differences in the gene expression patterns of the so-called “normal” surrounding tissues of AFTNs and CTNs (134). Principal component analysis and hierarchical clustering showed a distinct separation of the two subgroups, and statistical analysis revealed a significantly changed expression pattern of more than 300 genes (FDR < 1%). Among them, a decreased expression of thyroid-specific genes such as *TSHR*, *TPO*, and *PAX8* in the ST of AFTNs could be shown. Moreover, in the ST of AFTNs, a correlation between the presurgical TSH levels and the gene expression of *DIO1*, *TPO*, *SIAT1*, and the G<sub>i</sub>α-protein 1 (*GNAI1*) could be observed ( $r > 0.7$ ;  $P < 0.01$ ). These data show strong differences in the gene expression patterns of so-called normal thyroid tissues, which can be attributed to their different TSH exposure. Therefore, a careful selection of the normal reference tissue by presurgical assessment of the TSH levels and scintigraphy in iodine-deficient areas should be performed in studies investigating the gene expression patterns of thyroid tumors.

#### IV. Future Developments

##### A. Diagnostic implications

The primary goal of the evaluation of patients with nodular thyroid disease is the exclusion of thyroid malignancy. However, the identification and differentiation of rarely occurring thyroid cancer from frequently occurring benign nodular thyroid disease is challenging. Although FNA cytology (FNAC) represents the most sensitive and specific tool for the differential diagnosis of thyroid malignancy, there are important limitations. Although 75% of FNACs reveal a benign lesion and 5% a malignant lesion, up to 20% of FNACs reveal follicular neoplasia for which surgery is the only method to differentiate between follicular adenoma, follicular carcinoma, and the follicular variant of papillary carcinoma (135). Therefore, especially in the case of FNAC showing follicular neoplasia, markers would be helpful to discriminate in particular between follicular carcinoma and follicular adenoma. Up to now, many markers have been screened and evaluated, *e.g.*, *LGALS3*, hemoglobin, epsilon 1 (*HBE1*), keratin 19 (*CK19*), but also *TPO*, as well as genes related to initiation of transformation, like *PAX8/PPARγ* rearrangements and *BRAF* mutations (136–139). Although none of them are applicable to routine diagnostics so far, there is progress in molecular testing of fine-needle aspirates that may improve preoperative diagnosis. Most promising

techniques are based on the detection of tumor-specific mutations (*e.g.*, *BRAF* mutations, *RET/PTC* rearrangements) in the FNA material (140–143). These methods are characterized by a high specificity and easy detection of the mutations in genomic DNA. However, because not all FTCs and PTCs are characterized by the currently known mutations in *BRAF* or *RET/PTC* rearrangements or *PAX8/PPARγ* rearrangements, these methods are limited. Thus, the nearly genome-wide microarray investigations of both benign and malignant thyroid tumors and the comparison of these different studies may help to identify additional potentially useful markers. Although data from microarray investigations suggest that this technology has the sensitivity and specificity needed for a screening tool and as an adjunct to clinical pathological diagnosis in the future (33, 39), currently the amount and the quality of the RNA available from FNAs is limiting such an analysis in routine diagnosis. Therefore, an assay based on a limited number of differentiating genes, identified by sophisticated algorithms in comparative studies, appears to be more promising for FNAB application (30, 44). However, in contrast to the analysis of tumor-specific mutations, this approach of quantitatively measuring RNA markers is more susceptible to potential limitations of FNA such as limited and variable numbers of follicular cells obtained in each biopsy and the potential contamination by other cell types such as activated lymphocytes in patients with lymphocytic thyroiditis. Therefore, a correction for mRNA yield (*e.g.*, by measuring a housekeeping gene like  $\beta$ -actin) and thyroid specificity of mRNA extracted from a FNA sample (*e.g.*, by measuring a thyroid-specific gene like *TG*) seems mandatory. Nevertheless, the identification of genes or gene groups that correlate with nodular growth, thyroid dedifferentiation, or malignancy would make it possible to improve both diagnostics and therapy of thyroid nodular disease.

##### B. Genome analysis (SNP arrays, arrayCGH)

Up to now, the use of microarrays has mainly focused on the investigation of the transcriptome. However, advances in the microarray technology opened a wide range of new applications. The introduction of the Affymetrix GeneChip Mapping Arrays enables genotyping of up to 500,000 single nucleotide polymorphisms (SNPs) with an average distance between SNPs of 5.8 kb, which makes it possible to conduct large-scale linkage analyses, association studies, and copy number studies. This methodological development will also influence the genetic analysis of thyroid pathologies. In addition to a lack of iodine as the most prevalent factor for goiter development, familial clustering of goiters and the female predominance of goiters suggest a genetic background (144, 145). Actually, linkage analysis revealed two candidate regions, the multinodular goiter 1 locus and the Xp22 locus (146–148). A genomewide scan to detect susceptibility loci that predispose for euthyroid goiter using 450 microsatellite markers indicated genetic heterogeneity (149, 150). Therefore, linkage studies using SNP arrays will be an important adjunct to identify new susceptibility loci in an efficient way.

Comparative genomic hybridization (CGH) made a sig-

nificant impact on cancer cytogenetics. It provides the possibility to detect chromosomal copy number changes in cells, tissue samples, and formalin-fixed paraffin-embedded material. However, CGH to metaphase chromosomes provides only a limited resolution of 5–10 Mb for the detection of copy-number losses and gains, 2 Mb for amplifications, and the analysis requires a high level of cytogenetic expertise. The introduction of microarray-based CGH (arrayCGH) with a higher resolution could circumvent these limitations, thus making a broader application possible. Because classic CGH revealed several chromosomal imbalances in follicular thyroid adenoma, FTC, and the follicular variant of PTC (151–154), the investigation of thyroid tumors by microarray-based CGH might further resolve the pattern of these chromosomal aberrations or of the array expression patterns and might help to identify and specify the molecular etiology of thyroid tumors.

Additional new applications of microarray technology have been reported showing the high potential of this platform, *e.g.*, the combination of chromatin immunoprecipitation (ChIP) with hybridization of microarrays (ChIP-on-chip), which is a powerful way to explore sites of DNA-protein interaction across the whole genome, and the analysis of the methylation status of CpG islands in promoter regions. On the basis of these high resolution techniques, new genes involved in diseases as well as mechanistic aspects of diseases are likely to be discovered.

### Acknowledgments

Address all correspondence and requests for reprints to: R. Paschke, M.D., III. Medical Department, University of Leipzig, Ph.-Rosenthal-Str. 27, D-04103 Leipzig, Germany. E-mail: pasr@medizin.uni-leipzig.de

Disclosure Statement: M.E., K.K., A.K., and R.P. have nothing to declare. B.J. consults for Genzyme and received lecture fees from Novartis.

### References

1. Gershon D 2002 Microarray technology: an array of opportunities. *Nature* 416:885–891
2. Gress TM, Hoheisel JD, Lennon GG, Zehetner G, Lehrach H 1992 Hybridization fingerprinting of high-density cDNA-library arrays with cDNA pools derived from whole tissues. *Mamm Genome* 3:609–619
3. Schena M, Shalon D, Davis RW, Brown PO 1995 Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270:467–470
4. Fodor SP, Read JL, Pirrung MC, Stryer L, Lu AT, Solas D 1991 Light-directed, spatially addressable parallel chemical synthesis. *Science* 251:767–773
5. Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, Chee MS, Mittmann M, Wang C, Kobayashi M, Horton H, Brown EL 1996 Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol* 14:1675–1680
6. Chee M, Yang R, Hubbell E, Berno A, Huang XC, Stern D, Winkler J, Lockhart DJ, Morris MS, Fodor SP 1996 Accessing genetic information with high-density DNA arrays. *Science* 274:610–614
7. Wodicka L, Dong H, Mittmann M, Ho MH, Lockhart DJ 1997 Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat Biotechnol* 15:1359–1367
8. Eszlinger M, Krohn K, Paschke R 2001 Complementary DNA expression array analysis suggests a lower expression of signal transduction proteins and receptors in cold and hot thyroid nodules. *J Clin Endocrinol Metab* 86:4834–4842
9. Cheadle C, Vawter MP, Freed WJ, Becker KG 2003 Analysis of microarray data using Z score transformation. *J Mol Diagn* 5:73–81
10. Vawter MP, Barrett T, Cheadle C, Sokolov BP, Wood III WH, Donovan DM, Webster M, Freed WJ, Becker KG 2001 Application of cDNA microarrays to examine gene expression differences in schizophrenia. *Brain Res Bull* 55:641–650
11. van Staveren WC, Solis DW, Delys L, Venet D, Cappello M, Andry G, Dumont JE, Libert F, Detours V, Maenhaut C 2006 From the cover: gene expression in human thyrocytes and autonomous adenomas reveals suppression of negative feedbacks in tumorigenesis. *Proc Natl Acad Sci USA* 103:413–418
12. Hubbell E, Liu WM, Mei R 2002 Robust estimators for expression analysis. *Bioinformatics* 18:1585–1592
13. Li C, Wong WH 2001 Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc Natl Acad Sci USA* 98:31–36
14. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf TP 2003 Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Bio-statistics* 4:249–264
15. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP 2003 Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 31:e15
16. Wu Z, Irizarry RA, Gentleman R, Murillo FM, Spencer F 2003 A model-based background adjustment for oligonucleotide expression arrays. Technical report. Baltimore, MD: Johns Hopkins University, Department of Biostatistics Working Papers
17. Shedden K, Chen W, Kuick R, Ghosh D, Macdonald J, Cho KR, Giordano TJ, Gruber SB, Fearon ER, Taylor JM, Hanash S 2005 Comparison of seven methods for producing Affymetrix expression scores based on false discovery rates in disease profiling data. *BMC Bioinformatics* 6:26
18. Fuzarewicz K, Wiensch M 2003 Selecting differentially expressed genes for colon tumor classification. *Int J Applied Math Comp Science* 13:327–335
19. Fuzarewicz K, Kimmel M, Rzeszowska-Wolny J, Swierniak A 2003 A note on classification of gene expression data using support vector machines. *J Biol Systems* 11:43–56
20. Tibshirani R, Hastie T, Narasimhan B, Chu G 2002 Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci USA* 99:6567–6572
21. Westfall PH, Young SS 1993 Resampling-based multiple testing: examples and methods for multiple *P*-value adjustment. New York: John Wiley & Sons
22. Tusher VG, Tibshirani R, Chu G 2001 Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* 98:5116–5121
23. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES 1999 Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286:531–537
24. Hedenfalk I, Duggan D, Chen Y, Radmacher M, Bittner M, Simon R, Meltzer P, Gusterson B, Esteller M, Kallioniemi OP, Wilfond B, Borg A, Trent J 2001 Gene-expression profiles in hereditary breast cancer. *N Engl J Med* 344:539–548
25. Vapnik V 1995 The nature of statistical learning theory. New York: Springer-Verlag
26. Guyon I, Weston J, Barnhill S, Vapnik V 2002 Gene selection for cancer classification using support vector machines. *Machine Learning* 64:389–422
27. Aldred MA, Ginn-Pease ME, Morrison CD, Popkie AP, Gimm O, Hoang-Vu C, Krause U, Dralle H, Jhiang SM, Plass C, Eng C 2003 Caveolin-1 and caveolin-2, together with three bone morphogenetic protein-related genes, may encode novel tumor suppressors down-regulated in sporadic follicular thyroid carcinogenesis. *Cancer Res* 63:2864–2871
28. Giordano TJ, Au AY, Kuick R, Thomas DG, Rhodes DR, Wilhelm Jr KG, Vinco M, Misek DE, Sanders D, Zhu Z, Ciampi R, Hanash S, Chinnaiyan A, Clifton-Bligh RJ, Robinson BG, Nikiforov YE, Koenig RJ 2006 Delineation, functional validation, and bioinformatic

- matic evaluation of gene expression in thyroid follicular carcinomas with the PAX8-PPARG translocation. *Clin Cancer Res* 12:1983–1993
29. Barden CB, Shister KW, Zhu B, Guiter G, Greenblatt DY, Zeiger MA, Fahey III TJ 2003 Classification of follicular thyroid tumors by molecular signature: results of gene profiling. *Clin Cancer Res* 9:1792–1800
  30. Weber F, Shen L, Aldred MA, Morrison CD, Frilling A, Saji M, Schuppert F, Broelsch CE, Ringel MD, Eng C 2005 Genetic classification of benign and malignant thyroid follicular neoplasia based on a 3-gene combination. *J Clin Endocrinol Metab* 90:2512–2521
  31. Aldred MA, Huang Y, Liyanarachchi S, Pellegata NS, Gimm O, Jhiang S, Davuluri RV, de la Chapelle A, Eng C 2004 Papillary and follicular thyroid carcinomas show distinctly different microarray expression profiles and can be distinguished by a minimum of five genes. *J Clin Oncol* 22:3531–3539
  32. Chevillard S, Ugolin N, Vielh P, Ory K, Levalois C, Elliott D, Clayman GL, El Naggar AK 2004 Gene expression profiling of differentiated thyroid neoplasms: diagnostic and clinical implications. *Clin Cancer Res* 10:6586–6597
  33. Finley DJ, Zhu B, Barden CB, Fahey III TJ 2004 Discrimination of benign and malignant thyroid nodules by molecular profiling. *Ann Surg* 240:425–436
  34. Lacroix L, Lazar V, Michiels S, Ripoche H, Dessen P, Talbot M, Caillou B, Levillain JP, Schlumberger M, Bidart JM 2005 Follicular thyroid tumors with the PAX8-PPAR $\gamma$ 1 rearrangement display characteristic genetic alterations. *Am J Pathol* 167:223–231
  35. Lubitz CC, Gallagher LA, Finley DJ, Zhu B, Fahey III TJ 2005 Molecular analysis of minimally invasive follicular carcinomas by gene profiling. *Surgery* 138:1042–1049
  36. Lui WO, Foukakis T, Liden J, Thoppe SR, Dwight T, Hoog A, Zedenius J, Wallin G, Reimers M, Larsson C 2005 Expression profiling reveals a distinct transcription signature in follicular thyroid carcinomas with a PAX8-PPAR( $\gamma$ ) fusion oncogene. *Oncogene* 24:1467–1476
  37. Puskas LG, Juhasz F, Zarva A, Hackler Jr L, Farid NR 2005 Gene profiling identifies genes specific for well-differentiated epithelial thyroid tumors. *Cell Mol Biol (Noisy-le-grand)* 51:177–186
  38. Huang Y, Prasad M, Lemon WJ, Rempel H, Wright FA, Kornacker K, LiVolsi V, Frankel W, Kloos RT, Eng C, Pellegata NS, de la Chapelle A 2001 Gene expression in papillary thyroid carcinoma reveals highly consistent profiles. *Proc Natl Acad Sci USA* 98:15044–15049
  39. Lubitz CC, Ugras SK, Kazam JJ, Zhu B, Scognamiglio T, Chen YT, Fahey III TJ 2006 Microarray analysis of thyroid nodule fine-needle aspirates accurately classifies benign and malignant lesions. *J Mol Diagn* 8:490–498
  40. Finley DJ, Arora N, Zhu B, Gallagher L, Fahey III TJ 2004 Molecular profiling distinguishes papillary carcinoma from benign thyroid nodules. *J Clin Endocrinol Metab* 89:3214–3223
  41. Baris O, Savagner F, Nasser V, Loriod B, Granjeaud S, Guyétant S, Franc B, Rodien P, Rohmer V, Bertucci F, Birnbaum D, Malthiery Y, Reynier P, Houlgatte R 2004 Transcriptional profiling reveals coordinated up-regulation of oxidative metabolism genes in thyroid oncogenic tumors. *J Clin Endocrinol Metab* 89:994–1005
  42. Mazzanti C, Zeiger MA, Costouros NG, Umbrecht C, Westra WH, Smith D, Somervell H, Bevilacqua G, Alexander HR, Libutti SK, Costouros N 2004 Using gene expression profiling to differentiate benign versus malignant thyroid tumors. *Cancer Res* 64:2898–2903
  43. Frattini M, Ferrario C, Bressan P, Balestra D, De Cecco L, Mondellini P, Bongarzone I, Collini P, Gariboldi M, Pilotti S, Pierotti MA, Greco A 2004 Alternative mutations of BRAF, RET and NTRK1 are associated with similar but distinct gene expression patterns in papillary thyroid cancer. *Oncogene* 23:7436–7440
  44. Jarzab B, Wiench M, Fujarewicz K, Simek K, Jarzab M, Oczko-Wojciechowska M, Wloch J, Czarniecka A, Chmielik E, Lange D, Pawlaczek A, Szpak S, Gubala E, Swierniak A 2005 Gene expression profile of papillary thyroid cancer: sources of variability and diagnostic implications. *Cancer Res* 65:1587–1597
  45. Yano Y, Uematsu N, Yashiro T, Hara H, Ueno E, Miwa M, Tsujimoto G, Aiyoshi Y, Uchida K 2004 Gene expression profiling identifies platelet-derived growth factor as a diagnostic molecular marker for papillary thyroid carcinoma. *Clin Cancer Res* 10:2035–2043
  46. Giordano TJ, Kuick R, Thomas DG, Misek DE, Vinco M, Sanders D, Zhu Z, Ciampi R, Roh M, Shedden K, Gauger P, Doherty G, Thompson NW, Hanash S, Koenig RJ, Nikiforov YE 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
  47. Wreesmann VB, Sieczka EM, Socci ND, Hezel M, Belbin TJ, Childs G, Patel SG, Patel KN, Tallini G, Prystowsky M, Shaha AR, Kraus D, Shah JP, Rao PH, Ghossein R, Singh B 2004 Genome-wide profiling of papillary thyroid cancer identifies MUC1 as an independent prognostic marker. *Cancer Res* 64:3780–3789
  48. Eszlinger M, Wiench M, Jarzab B, Krohn K, Beck M, Lauter J, Gubala E, Fujarewicz K, Swierniak A, Paschke R 2006 Meta- and reanalysis of gene expression profiles of hot and cold thyroid nodules and papillary thyroid carcinoma for gene groups. *J Clin Endocrinol Metab* 91:1934–1942
  49. Stolf BS, Abreu CM, Mahler-Araujo MB, Dellamano M, Martins WK, de Carvalho MB, Curado MP, Diaz JP, Fabri A, Brentani H, Carvalho AF, Soares FA, Kowalski LP, Hirata Jr R, Reis LF 2005 Expression profile of malignant and non-malignant diseases of the thyroid gland reveals altered expression of a common set of genes in goiter and papillary carcinomas. *Cancer Lett* 227:59–73
  50. Baris O, Mirebeau-Prunier D, Savagner F, Rodien P, Ballester B, Loriod B, Granjeaud S, Guyétant S, Franc B, Houlgatte R, Reynier P, Malthiery Y 2005 Gene profiling reveals specific oncogenic mechanisms and signaling pathways in oncogenic and papillary thyroid carcinoma. *Oncogene* 24:4155–4161
  51. Eszlinger M, Krohn K, Beck M, Kipling D, Forbes-Robertson S, Lauter J, Tonjes A, Wynford-Thomas D, Paschke R 2006 Comparison of differential gene expression of hot and cold thyroid nodules with primary epithelial cell culture models by investigation of co-regulated gene sets. *Biochim Biophys Acta* 1763:263–271
  52. Eszlinger M, Krohn K, Frenzel R, Kropf S, Tonjes A, Paschke R 2004 Gene expression analysis reveals evidence for inactivation of the TGF- $\beta$  signaling cascade in autonomously functioning thyroid nodules. *Oncogene* 23:795–804
  53. Wattel S, Mircescu H, Venet D, Burniat A, Franc B, Frank S, Andry G, Van Sande J, Rocmans P, Dumont JE, Detours V, Maenhaut C 2005 Gene expression in thyroid autonomous adenomas provides insight into their physiopathology. *Oncogene* 24:6902–6916
  54. Eszlinger M, Krohn K, Berger K, Lauter J, Kropf S, Beck M, Fuhrer D, Paschke R 2005 Gene expression analysis reveals evidence for increased expression of cell cycle-associated genes and Gq-protein-protein kinase C signaling in cold thyroid nodules. *J Clin Endocrinol Metab* 90:1163–1170
  55. Mazzaferri EL, Samaan N 1993 *Endocrine tumors*. Cambridge, UK: Blackwell; 278
  56. Borrello MG, Alberti L, Fischer A, Degl'innocenti D, Ferrario C, Gariboldi M, Marchesi F, Allavena P, Greco A, Collini P, Pilotti S, Cassinelli G, Bressan P, Fugazzola L, Mantovani A, Pierotti MA 2005 Induction of a proinflammatory program in normal human thyrocytes by the RET/PTC1 oncogene. *Proc Natl Acad Sci USA* 102:14825–14830
  57. Detours V, Wattel S, Venet D, Hutsebaut N, Bogdanova T, Tronko MD, Dumont JE, Franc B, Thomas G, Maenhaut C 2005 Absence of a specific radiation signature in post-Chernobyl thyroid cancers. *Br J Cancer* 92:1545–1552
  58. Melillo RM, Castellone MD, Guarino V, De Falco V, Cirafici AM, Salvatore G, Caiazzo F, Basolo F, Giannini R, Kruhfoffer M, Orntoft T, Fusco A, Santoro M 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
  59. Hoang-Vu C, Dralle H, Scheumann G, Maenhaut C, Horn R, von zur Muhlen A, Brabant G 1992 Gene expression of differentiation- and dedifferentiation markers in normal and malignant human thyroid tissues. *Exp Clin Endocrinol* 100:51–56
  60. Loh KC, Greenspan FS, Dong F, Miller TR, Yeo PP 1999 Influence of lymphocytic thyroiditis on the prognostic outcome of patients with papillary thyroid carcinoma. *J Clin Endocrinol Metab* 84:458–463

61. Ciampi R, Knauf JA, Kerler R, Gandhi M, Zhu Z, Nikiforova MN, Rabes HM, Fagin JA, Nikiforov YE 2005 Oncogenic AKAP9-BRAF fusion is a novel mechanism of MAPK pathway activation in thyroid cancer. *J Clin Invest* 115:94–101
62. Kimura ET, Nikiforova MN, Zhu Z, Knauf JA, Nikiforov YE, Fagin JA 2003 High prevalence of BRAF mutations in thyroid cancer: genetic evidence for constitutive activation of the RET/PTC-RAS-BRAF signaling pathway in papillary thyroid carcinoma. *Cancer Res* 63:1454–1457
63. Soares P, Trovisco V, Rocha AS, Lima J, Castro P, Preto A, Maximo V, Botelho T, Seruca R, Sobrinho-Simoes M 2003 BRAF mutations and RET/PTC rearrangements are alternative events in the etiopathogenesis of PTC. *Oncogene* 22:4578–4580
64. Nikiforova MN, Kimura ET, Gandhi M, Biddinger PW, Knauf JA, Basolo F, Zhu Z, Giannini R, Salvatore G, Fusco A, Santoro M, Fagin JA, Nikiforov YE 2003 BRAF mutations in thyroid tumors are restricted to papillary carcinomas and anaplastic or poorly differentiated carcinomas arising from papillary carcinomas. *J Clin Endocrinol Metab* 88:5399–5404
65. Puxeddu E, Moretti S, Elisei R, Romei C, Pascucci R, Martinelli M, Marino C, Avenia N, Rossi ED, Fadda G, Cavaliere A, Ribacchi R, Falorni A, Pontecorvi A, Pacini F, Pinchera A, Santeusano F 2004 BRAF(V599E) mutation is the leading genetic event in adult sporadic papillary thyroid carcinomas. *J Clin Endocrinol Metab* 89:2414–2420
66. Liu RT, Chen YJ, Chou FF, Li CL, Wu WL, Tsai PC, Huang CC, Cheng JT 2005 No correlation between BRAFV600E mutation and clinicopathological features of papillary thyroid carcinomas in Taiwan. *Clin Endocrinol (Oxf)* 63:461–466
67. Fugazzola L, Mannavola D, Cirello V, Vannucchi G, Muzza M, Vicentini L, Beck-Peccoz P 2004 BRAF mutations in an Italian cohort of thyroid cancers. *Clin Endocrinol (Oxf)* 61:239–243
68. Kim TY, Kim WB, Song JY, Rhee YS, Gong G, Cho YM, Kim SY, Kim SC, Hong SJ, Shong YK 2005 The BRAFV600E mutation is not associated with poor prognostic factors in Korean patients with conventional papillary thyroid microcarcinoma. *Clin Endocrinol (Oxf)* 63:588–593
69. Adeniran AJ, Zhu Z, Gandhi M, Steward DL, Fidler JP, Giordano TJ, Biddinger PW, Nikiforov YE 2006 Correlation between genetic alterations and microscopic features, clinical manifestations, and prognostic characteristics of thyroid papillary carcinomas. *Am J Surg Pathol* 30:216–222
70. Namba H, Nakashima M, Hayashi T, Hayashida N, Maeda S, Rogounovitch TI, Ohtsuru A, Saenko VA, Kanematsu T, Yamashita S 2003 Clinical implication of hot spot BRAF mutation, V599E, in papillary thyroid cancers. *J Clin Endocrinol Metab* 88:4393–4397
71. Cerutti JM, Delcelo R, Amadei MJ, Nakabashi C, Maciel RM, Peterson B, Shoemaker J, Riggins GJ 2004 A preoperative diagnostic test that distinguishes benign from malignant thyroid carcinoma based on gene expression. *J Clin Invest* 113:1234–1242
72. Fagin JA 2002 Minireview: branded from the start—distinct oncogenic initiating events may determine tumor fate in the thyroid. *Mol Endocrinol* 16:903–911
73. Kroll TG, Sarraf P, Pecciarini L, Chen CJ, Mueller E, Spiegelman BM, Fletcher JA 2000 PAX8-PPAR $\gamma$ 1 fusion oncogene in human thyroid carcinoma [corrected]. *Science [Erratum (2000) 289:1474]* 289:1357–1360
74. Marques AR, Espadinha C, Catarino AL, Moniz S, Pereira T, Sobrinho LG, Leite V 2002 Expression of PAX8-PPAR $\gamma$  1 rearrangements in both follicular thyroid carcinomas and adenomas. *J Clin Endocrinol Metab* 87:3947–3952
75. Nikiforova MN, Biddinger PW, Caudill CM, Kroll TG, Nikiforov YE 2002 PAX8-PPAR $\gamma$  rearrangement in thyroid tumors: RT-PCR and immunohistochemical analyses. *Am J Surg Pathol* 26:1016–1023
76. Cheung L, Messina M, Gill A, Clarkson A, Learoyd D, Delbridge L, Wentworth J, Phillips J, Clifton-Bligh R, Robinson BG 2003 Detection of the PAX8-PPAR $\gamma$  fusion oncogene in both follicular thyroid carcinomas and adenomas. *J Clin Endocrinol Metab* 88:354–357
77. Gregory PJ, Wang X, Allard BL, Sahin M, Wang XL, Hay ID, Hiddinga HJ, Deshpande SS, Kroll TG, Grebe SK, Eberhardt NL, McIver B 2004 The PAX8/PPAR $\gamma$  fusion oncoprotein transforms immortalized human thyrocytes through a mechanism probably involving wild-type PPAR $\gamma$  inhibition. *Oncogene* 23:3634–3641
78. French CA, Alexander EK, Cibas ES, Nose V, Laguette J, Faquin W, Garber J, Moore Jr F, Fletcher JA, Larsen PR, Kroll TG 2003 Genetic and biological subgroups of low-stage follicular thyroid cancer. *Am J Pathol* 162:1053–1060
79. Nikiforova MN, Lynch RA, Biddinger PW, Alexander EK, Dorn GW, Tallini G, Kroll TG, Nikiforov YE 2003 RAS point mutations and PAX8-PPAR $\gamma$  rearrangement in thyroid tumors: evidence for distinct molecular pathways in thyroid follicular carcinoma. *J Clin Endocrinol Metab* 88:2318–2326
80. Eszlinger M, Krohn K, Kratzsch J, Voigt C, Paschke R 2001 Growth factor expression in cold and hot thyroid nodules. *Thyroid* 11:125–135
81. Gärtner R, Schopohl D, Schaefer S, Dugrillon A, Erdmann A, Toda S, Bechtner G 1997 Regulation of transforming growth factor  $\beta$ 1 messenger ribonucleic acid expression in porcine thyroid follicles in vitro by growth factors, iodine, or  $\delta$ -iodolactone. *Thyroid* 7:633–640
82. Voigt C, Holzapfel H, Paschke R 2000 Expression of  $\beta$ -arrestins in toxic and cold thyroid nodules. *FEBS Lett* 486:208–212
83. Chen W, Kirkbride KC, How T, Nelson CD, Mo J, Frederick JP, Wang XF, Lefkowitz RJ, Blobel GC 2003 Beta-arrestin 2 mediates endocytosis of type III TGF- $\beta$  receptor and down-regulation of its signaling. *Science* 301:1394–1397
84. Frenzel R, Krohn K, Eszlinger M, Tonjes A, Paschke R 2005 Sialylation of human thyrotropin receptor improves and prolongs its cell-surface expression. *Mol Pharmacol* 68:1106–1113
85. Persani L, Lania A, Alberti L, Romoli R, Mantovani G, Filetti S, Spada A, Conti M 2000 Induction of specific phosphodiesterase isoforms by constitutive activation of the cAMP pathway in autonomous thyroid adenomas. *J Clin Endocrinol Metab* 85:2872–2878
86. Voigt C, Holzapfel HP, Meyer S, Paschke R 2004 Increased expression of G-protein-coupled receptor kinases 3 and 4 in hyperfunctioning thyroid nodules. *J Endocrinol* 182:173–182
87. Eszlinger M, Holzapfel HP, Voigt C, Arkenau C, Paschke R 2004 RGS 2 expression is regulated by TSH and inhibits TSH receptor signaling. *Eur J Endocrinol* 151:383–390
88. Tonjes A, Miedlich S, Holzapfel HP, Eszlinger M, Arkenau C, Paschke R 2004 Expression of regulators of G protein signaling mRNA is differentially regulated in hot and cold thyroid nodules. *Thyroid* 14:896–901
89. Uyttersprot N, Costagliola S, Dumont JE, Miot F 1999 Requirement for cAMP-response element (CRE) binding protein/CRE modulator transcription factors in thyrotropin-induced proliferation of dog thyroid cells in primary culture. *Eur J Biochem* 259:370–378
90. Finley DJ, Lubitz CC, Wei C, Zhu B, Fahey III TJ 2005 Advancing the molecular diagnosis of thyroid nodules: defining benign lesions by molecular profiling. *Thyroid* 15:562–568
91. Krohn K, Fuhrer D, Bayer Y, Eszlinger M, Brauer V, Neumann S, Paschke R 2005 Molecular pathogenesis of euthyroid and toxic multinodular goiter. *Endocr Rev* 26:504–524
92. Krohn K, Emmrich P, Ott N, Paschke R 1999 Increased thyroid epithelial cell proliferation in toxic thyroid nodules. *Thyroid* 9:241–246
93. Krohn K, Stricker I, Emmrich P, Paschke R 2003 Cold thyroid nodules show a marked increase in proliferation markers. *Thyroid* 13:569–575
94. Esapa CT, Johnson SJ, Kendall-Taylor P, Lennard TW, Harris PE 1999 Prevalence of Ras mutations in thyroid neoplasia. *Clin Endocrinol (Oxf)* 50:529–535
95. Krohn K, Reske A, Ackermann F, Muller A, Paschke R 2001 Ras mutations are rare in solitary cold and toxic thyroid nodules. *Clin Endocrinol (Oxf)* 55:241–248
96. Gire V, Wynford-Thomas D 2000 RAS oncogene activation induces proliferation in normal human thyroid epithelial cells without loss of differentiation. *Oncogene* 19:737–744
97. Bachrach LK, Eggo MC, Mak WW, Burrow GN 1985 Phorbol esters stimulate growth and inhibit differentiation in cultured thyroid cells. *Endocrinology* 116:1603–1609

98. Haye B, Aublin JL, Champion S, Lambert B, Jacquemin C 1985 Tetradecanoyl phorbol-13-acetate counteracts the responsiveness of cultured thyroid cells to thyrotropin. *Biochem Pharmacol* 34: 3795–3802
99. Roger PP, Reuse S, Servais P, Van Heuverswyn B, Dumont JE 1986 Stimulation of cell proliferation and inhibition of differentiation expression by tumor-promoting phorbol esters in dog thyroid cells in primary culture. *Cancer Res* 46:898–906
100. Takasu N, Shimizu Y, Yamada T 1987 Tumour promoter 12-O-tetradecanoylphorbol 13-acetate and epidermal growth factor stimulate proliferation and inhibit differentiation of porcine thyroid cells in primary culture. *J Endocrinol* 113:485–487
101. Vassart G, Dumont JE 1992 The thyrotropin receptor and the regulation of thyrocyte function and growth. *Endocr Rev* 13:596–611
102. Van Herle AJ, Agatep ML, Padua III DN, Totanes TL, Canlapan DV, Van Herle HM, Juillard GJ 1990 Effects of 13 *cis*-retinoic acid on growth and differentiation of human follicular carcinoma cells (UCLA R0 82 W-1) *in vitro*. *J Clin Endocrinol Metab* 71:755–763
103. Schmutzler C, Brtko J, Bienert K, Kohrle J 1996 Effects of retinoids and role of retinoic acid receptors in human thyroid carcinomas and cell lines derived therefrom. *Exp Clin Endocrinol Diabetes* 104(Suppl 4):16–19
104. Schreck R, Schnieders F, Schmutzler C, Kohrle J 1994 Retinoids stimulate type I iodothyronine 5'-deiodinase activity in human follicular thyroid carcinoma cell lines. *J Clin Endocrinol Metab* 79:791–798
105. Donovan M, Olofsson B, Gustafson AL, Dencker L, Eriksson U 1995 The cellular retinoic acid-binding proteins. *J Steroid Biochem Mol Biol* 53:459–465
106. Aogi K, Kitahara K, Buley I, Backdahl M, Tahara H, Sugino T, Tarin D, Goodison S 1998 Telomerase activity in lesions of the thyroid: application to diagnosis of clinical samples including fine-needle aspirates. *Clin Cancer Res* 4:1965–1970
107. Bernet VJ, Anderson J, Vaishnav Y, Solomon B, Adair CF, Saji M, Burman KD, Burch HB, Ringel MD 2002 Determination of galectin-3 messenger ribonucleic acid overexpression in papillary thyroid cancer by quantitative reverse transcription-polymerase chain reaction. *J Clin Endocrinol Metab* 87:4792–4796
108. Casey MB, Lohse CM, Lloyd RV 2003 Distinction between papillary thyroid hyperplasia and papillary thyroid carcinoma by immunohistochemical staining for cytokeratin 19, galectin-3, and HBME-1. *Endocr Pathol* 14:55–60
109. Haugen BR, Nawaz S, Markham N, Hashizumi T, Shroyer AL, Werness B, Shroyer KR 1997 Telomerase activity in benign and malignant thyroid tumors. *Thyroid* 7:337–342
110. Inohara H, Honjo Y, Yoshii T, Akahani S, Yoshida J, Hattori K, Okamoto S, Sawada T, Raz A, Kubo T 1999 Expression of galectin-3 in fine-needle aspirates as a diagnostic marker differentiating benign from malignant thyroid neoplasms. *Cancer* 85:2475–2484
111. Ippolito A, Vella V, La Rosa GL, Pellegriti G, Vigneri R, Belfiore A 2001 Immunostaining for Met/HGF receptor may be useful to identify malignancies in thyroid lesions classified suspicious at fine-needle aspiration biopsy. *Thyroid* 11:783–787
112. Mase T, Funahashi H, Koshikawa T, Imai T, Nara Y, Tanaka Y, Nakao A 2003 HBME-1 immunostaining in thyroid tumors especially in follicular neoplasm. *Endocr J* 50:173–177
113. Raphael SJ 2002 The meanings of markers: ancillary techniques in diagnosis of thyroid neoplasia. *Endocr Pathol* 13:301–311
114. Sack MJ, Astengo-Osuna C, Lin BT, Battifora H, LiVolsi VA 1997 HBME-1 immunostaining in thyroid fine-needle aspirations: a useful marker in the diagnosis of carcinoma. *Mod Pathol* 10:668–674
115. Saggiorato E, Cappia S, De Giuli P, Mussa A, Pancani G, Caraci P, Angeli A, Orlandi F 2001 Galectin-3 as a presurgical immunocytochemical marker of minimally invasive follicular thyroid carcinoma. *J Clin Endocrinol Metab* 86:5152–5158
116. Kholova I, Ryska A, Ludvikova M, Cap J, Pecan L 2003 Dipeptidyl peptidase IV expression in thyroid cytology: retrospective histologically confirmed study. *Cytopathology* 14:27–31
117. Giannini R, Faviana P, Cavinato T, Elisei R, Pacini F, Berti P, Fontanini G, Ugolini C, Camacci T, De Ieso K, Miccoli P, Pinchera A, Basolo F 2003 Galectin-3 and oncofetal-fibronectin expression in thyroid neoplasia as assessed by reverse transcription-polymerase chain reaction and immunochemistry in cytologic and pathologic specimens. *Thyroid* 13:765–770
118. Papotti M, Rodriguez J, De Pompa R, Bartolazzi A, Rosai J 2005 Galectin-3 and HBME-1 expression in well-differentiated thyroid tumors with follicular architecture of uncertain malignant potential. *Mod Pathol* 18:541–546
119. Prasad ML, Huang Y, Pellegata NS, de la Chapelle A, Kloos RT 2004 Hashimoto's thyroiditis with papillary thyroid carcinoma (PTC)-like nuclear alterations express molecular markers of PTC. *Histopathology* 45:39–46
120. Saggiorato E, De Pompa R, Volante M, Cappia S, Arecco F, Dei Tos AP, Orlandi F, Papotti M 2005 Characterization of thyroid 'follicular neoplasms' in fine-needle aspiration cytological specimens using a panel of immunohistochemical markers: a proposal for clinical application. *Endocr Relat Cancer* 12:305–317
121. Takano T, Miyauchi A, Matsuzuka F, Kuma K, Amino N 2000 Expression of oncofetal fibronectin messenger ribonucleic acid in fibroblasts in the thyroid: a possible cause of false positive results in molecular-based diagnosis of thyroid carcinomas. *J Clin Endocrinol Metab* 85:765–768
122. Grutzmann R, Boriss H, Ammerpohl O, Luttes J, Kalthoff H, Schackert HK, Kloppel G, Saeger HD, Pilarsky C 2005 Meta-analysis of microarray data on pancreatic cancer defines a set of commonly dysregulated genes. *Oncogene* 24:5079–5088
123. Kebebew E, Peng M, Reiff E, McMillan A 2006 Diagnostic and extent of disease multigene assay for malignant thyroid neoplasms. *Cancer* 106:2592–2597
124. Venet D, Pécasse F, Maenhaut C, Bersini H 2001 Separation of samples into their constituents using gene expression data. *Bioinformatics* 17(Suppl 1):S279–S287
125. Lubitz CC, Fahey III TJ 2006 Gene expression profiling of thyroid tumors—clinical applicability. *Nat Clin Pract Endocrinol Metab* 2:472–473
126. Tan PK, Downey TJ, Spitznagel Jr EL, Xu P, Fu D, Dimitrov DS, Lempicki RA, Raaka BM, Cam MC 2003 Evaluation of gene expression measurements from commercial microarray platforms. *Nucleic Acids Res* 31:5676–5684
127. Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, Spellman P, Iyer V, Jeffrey SS, Van de Rijn M, Waltham M, Pergamenschikov A, Lee JC, Lashkari D, Shalon D, Myers TG, Weinstein JN, Botstein D, Brown PO 2000 Systematic variation in gene expression patterns in human cancer cell lines. *Nat Genet* 24:227–235
128. Kuo WP, Jansen TK, Butte AJ, Ohno-Machado L, Kohane IS 2002 Analysis of matched mRNA measurements from two different microarray technologies. *Bioinformatics* 18:405–412
129. Larkin JE, Frank BC, Gavras H, Sultana R, Quackenbush J 2005 Independence and reproducibility across microarray platforms. *Nat Methods* 2:337–344
130. Bammler T, Beyer RP, Bhattacharya S, Boorman GA, Boyles A, Bradford BU, Bumgarner RE, Bushel PR, Chaturvedi K, Choi D, Cunningham ML, Deng S, Dressman HK, Fannin RD, Farin FM, Freedman JH, Fry RC, Harper A, Humble MC, Hurban P, Kavanagh TJ, Kaufmann WK, Kerr KF, Jing L, Lapidus JA, Lasarev MR, Li J, Li YJ, Lobenhofer EK, Lu X, Malek RL, Milton S, Nagalla SR, O'Malley JP, Palmer VS, Pattee P, Paules RS, Perou CM, Phillips K, Qin LX, Qiu Y, Quigley SD, Rodland M, Rusyn I, Samson LD, Schwartz DA, Shi Y, Shin JL, Sieber SO, Sliker S, Speer MC, Spencer PS, Sproles DI, Swenberg JA, Suk WA, Sullivan RC, Tian R, Tennant RW, Todd SA, Tucker CJ, Van Houten B, Weis BK, Xuan S, Zarbl H 2005 Standardizing global gene expression analysis between laboratories and across platforms. *Nat Methods* 2:351–356
131. Irizarry RA, Warren D, Spencer F, Kim IF, Biswal S, Frank BC, Gabrielson E, Garcia JG, Geoghegan J, Grumkin G, Griffin C, Hilder S, Hoffman E, Jedlicka AE, Kawasaki E, Martinez-Murillo F, Morsberger L, Lee H, Petersen D, Quackenbush J, Scott A, Wilson M, Yang Y, Ye SQ, Yu W 2005 Multiple-laboratory comparison of microarray platforms. *Nat Methods* 2:345–350
132. Griffith OL, Melck A, Jones SJ, Wiseman SM 2006 Meta-analysis and meta-review of thyroid cancer gene expression profiling studies identifies important diagnostic biomarkers. *J Clin Oncol* 24: 5043–5051
133. Bruno R, Ferretti E, Tosi E, Arturi F, Giannasio P, Mattei T,



- Scipioni A, Presta I, Morisi R, Gulino A, Filetti S, Russo D 2005 Modulation of thyroid-specific gene expression in normal and nodular human thyroid tissues from adults: an *in vivo* effect of thyrotropin. *J Clin Endocrinol Metab* 90:5692–5697
134. Eszlinger M, Krohn K, Toenjes A, Läuter J, Paschke R 2007 Variability of gene expression of “normal” thyroid tissue. *Exp Clin Endocrinol Diabetes* 115(Suppl 1):8 (Abstract)
135. Gharib H 1997 Changing concepts in the diagnosis and management of thyroid nodules. *Endocrinol Metab Clin North Am* 26:777–800
136. Bartolazzi A, Gasbarri A, Papotti M, Bussolati G, Lucante T, Khan A, Inohara H, Marandino F, Orlandi F, Nardi F, Vecchione A, Tecce R, Larsson O 2001 Application of an immunodiagnostic method for improving preoperative diagnosis of nodular thyroid lesions. *Lancet* 357:1644–1650
137. de Micco C, Ruf J, Chrestian MA, Gros N, Henry JF, Carayon P 1991 Immunohistochemical study of thyroid peroxidase in normal, hyperplastic, and neoplastic human thyroid tissues. *Cancer* 67:3036–3041
138. Fagin JA 2002 Perspective: lessons learned from molecular genetic studies of thyroid cancer—insights into pathogenesis and tumor-specific therapeutic targets. *Endocrinology* 143:2025–2028
139. Gimm O 2001 Thyroid cancer. *Cancer Lett* 163:143–156
140. Cohen Y, Rosenbaum E, Clark DP, Zeiger MA, Umbricht CB, Tufano RP, Sidransky D, Westra WH 2004 Mutational analysis of BRAF in fine needle aspiration biopsies of the thyroid: a potential application for the preoperative assessment of thyroid nodules. *Clin Cancer Res* 10:2761–2765
141. Jin L, Sebo TJ, Nakamura N, Qian X, Oliveira A, Majerus JA, Johnson MR, Lloyd RV 2006 BRAF mutation analysis in fine needle aspiration (FNA) cytology of the thyroid. *Diagn Mol Pathol* 15:136–143
142. Rowe LM, Bentz BM, Bentz JM 2006 Utility of BRAF V600E mutation detection in cytologically indeterminate thyroid nodules. *Cytojournal* 3:10
143. Salvatore G, Giannini R, Faviana P, Caleo A, Migliaccio I, Fagin JA, Nikiforov YE, Troncone G, Palombini L, Basolo F, Santoro M 2004 Analysis of BRAF point mutation and RET/PTC rearrangement refines the fine-needle aspiration diagnosis of papillary thyroid carcinoma. *J Clin Endocrinol Metab* 89:5175–5180
144. Malamos B, Koutras DA, Kostamis P, Rigopoulos GA, Zerefos NS, Yataganas XA 1967 Endemic goitre in Greece: a study of 379 twin pairs. *J Med Genet* 4:16–18
145. Hadjidakis SG, Koutras DA, Daikos GK 1964 Endemic goitre in Greece: family studies. *J Med Genet* 38:82–87
146. Bignell GR, Canzian F, Shayeghi M, Stark M, Shugart YY, Biggs P, Mangion J, Hamoudi R, Rosenblatt J, Buu P, Sun S, Stoffer SS, Goldgar DE, Romeo G, Houlston RS, Narod SA, Stratton MR, Foulkes WD 1997 Familial nontoxic multinodular thyroid goiter locus maps to chromosome 14q but does not account for familial nonmedullary thyroid cancer. *Am J Hum Genet* 61:1123–1130
147. Capon F, Tacconelli A, Giardina E, Sciacchitano S, Bruno R, Tassi V, Trischitta V, Filetti S, Dallapiccola B, Novelli G 2000 Mapping a dominant form of multinodular goiter to chromosome Xp22. *Am J Hum Genet* 67:1004–1007
148. Neumann S, Willgerodt H, Ackermann F, Reske A, Jung M, Reis A, Paschke R 1999 Linkage of familial euthyroid goiter to the multinodular goiter-1 locus and exclusion of the candidate genes thyroglobulin, thyroperoxidase, and Na<sup>+</sup>/I<sup>-</sup> symporter. *J Clin Endocrinol Metab* 84:3750–3756
149. Bayer Y, Neumann S, Meyer B, Ruschendorf F, Reske A, Brix T, Hegedus L, Langer P, Nurnberg P, Paschke R 2004 Genome-wide linkage analysis reveals evidence for four new susceptibility loci for familial euthyroid goiter. *J Clin Endocrinol Metab* 89:4044–4052
150. Bottcher Y, Eszlinger M, Tonjes A, Paschke R 2005 The genetics of euthyroid familial goiter. *Trends Endocrinol Metab* 16:314–319
151. Rodrigues RF, Roque L, Rosa-Santos J, Cid O, Soares J 2004 Chromosomal imbalances associated with anaplastic transformation of follicular thyroid carcinomas. *Br J Cancer* 90:492–496
152. Roque L, Rodrigues R, Pinto A, Moura-Nunes V, Soares J 2003 Chromosome imbalances in thyroid follicular neoplasms: a comparison between follicular adenomas and carcinomas. *Genes Chromosomes Cancer* 36:292–302
153. Castro P, Eknaes M, Teixeira MR, Danielsen HE, Soares P, Lothe RA, Sobrinho-Simoes M 2005 Adenomas and follicular carcinomas of the thyroid display two major patterns of chromosomal changes. *J Pathol* 206:305–311
154. Wreesmann VB, Ghossein RA, Hezel M, Banerjee D, Shaha AR, Tuttle RM, Shah JP, Rao PH, Singh B 2004 Follicular variant of papillary thyroid carcinoma: genome-wide appraisal of a controversial entity. *Genes Chromosomes Cancer* 40:355–364
155. Wasenius VM, Hemmer S, Kettunen E, Knuutila S, Franssila K, Joensuu H 2003 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* 9:68–75
156. Prasad ML, Pellegata NS, Kloos RT, Barbacioru C, Huang Y, de la Chapelle A 2004 CITED1 protein expression suggests papillary thyroid carcinoma in high throughput tissue microarray-based study. *Thyroid* 14:169–175
157. Zou M, Famulski KS, Parhar RS, Baitei E, Al Mohanna FA, Farid NR, Shi Y 2004 Microarray analysis of metastasis-associated gene expression profiling in a murine model of thyroid carcinoma pulmonary metastasis: identification of S100A4 (Mts1) gene overexpression as a poor prognostic marker for thyroid carcinoma. *J Clin Endocrinol Metab* 89:6146–6154
158. Chen KT, Lin JD, Chao TC, Hsueh C, Chang CA, Weng HF, Chan EC 2001 Identifying differentially expressed genes associated with metastasis of follicular thyroid cancer by cDNA expression array. *Thyroid* 11:41–46
159. Hamada A, Mankovskaya S, Saenko V, Rogounovitch T, Mine M, Namba H, Nakashima M, Demidchik Y, Demidchik E, Yamashita S 2005 Diagnostic usefulness of PCR profiling of the differentially expressed marker genes in thyroid papillary carcinomas. *Cancer Lett* 224:289–301
160. Takano T, Hasegawa Y, Miyauchi A, Matsuzuka F, Yoshida H, Kuma K, Hayashi N, Nakamori S, Amino N 2002 Quantitative analysis of osteonectin mRNA in thyroid carcinomas. *Endocr J* 49:511–516
161. Takano T, Miyauchi A, Yoshida H, Kuma K, Amino N 2004 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* 90:1600–1605
162. Jacques C, Baris O, Prunier-Mirebeau D, Savagner F, Rodien P, Rohmer V, Franc B, Guyetant S, Malthiery Y, Reynier P 2005 Two-step differential expression analysis reveals a new set of genes involved in thyroid oncogenic tumors. *J Clin Endocrinol Metab* 90:2314–2320
163. Haugen BR, Larson LL, Pugazhenth U, Hays WR, Klopper JP, Kramer CA, Sharma V 2004 Retinoic acid and retinoid X receptors are differentially expressed in thyroid cancer and thyroid carcinoma cell lines and predict response to treatment with retinoids. *J Clin Endocrinol Metab* 89:272–280
164. Aksoy M, Giles Y, Kapran Y, Terzioglu T, Tezelman S 2005 Expression of bcl-2 in papillary thyroid cancers and its prognostic value. *Acta Chir Belg* 105:644–648